DOCUMENT RESUME

ED 272 534 TM 860 450

AUTHOR	Silverman, Leslie J., Ed.; Taeuber, Richard C., Ed.
TITLE	Invited Papers: Elementary/Secondary Education Data Redesign Project, October 1985.
SPONS AGENCY	National Center for Education Statistics (ED), Washington, DC.
PUB DATE	Oct 85
NOTE	798p.; For a synthesis of the papers, a comparison volume, and individual papers, see TM 860 491-492 and TM 860 451-490.
PUB TYPE	Collected Works - General (020) -- Information Analyses (070) -- Viewpoints (120)
EDRS PRICE	MF05/PC32 Plus Postage.
DESCRIPTORS	*Data Collection; *Educational Assessment; Educational Needs; *Educational Policy; Elementary Secondary Education; *National Programs; National Surveys; Program Evaluation; Research Design; *Research Needs; *School Statistics
IDENTIFIERS	*National Center for Educational Statistics

ABSTRACT

This document includes invited papers written for the National Center for Education Statistics (NCES) Elementary/Secondary Education Data Redesign Project. Early in 1985, NCES began a thorough review of its data collection program. The purpose of the review was to address questions of the suitability, scope, reliability, and timeliness of statistical data in terms of the purposes those data serve, relative to instructional and administrative needs and relative to education policy issues. To initiate the dialogue, a number of individuals and organizations were invited to prepare a paper following guidelines provided by NCES. Individuals invited were likely to (1) use quantitative approaches in their analyses, and thus be data users, and (2) have the interest and time to produce a "thought paper," not a research paper. There was no intent to represent known positions or to provide balanced representations across sectors of the education community. Organizations were selected because of an identification with education issues and concerns. The results of these efforts may well be the design of a new plan for national data collections from institutions and individuals to be implemented over several years. A total of 38 individual papers, 16 organization papers, and 3 agency papers are included in this volume. (LMO)

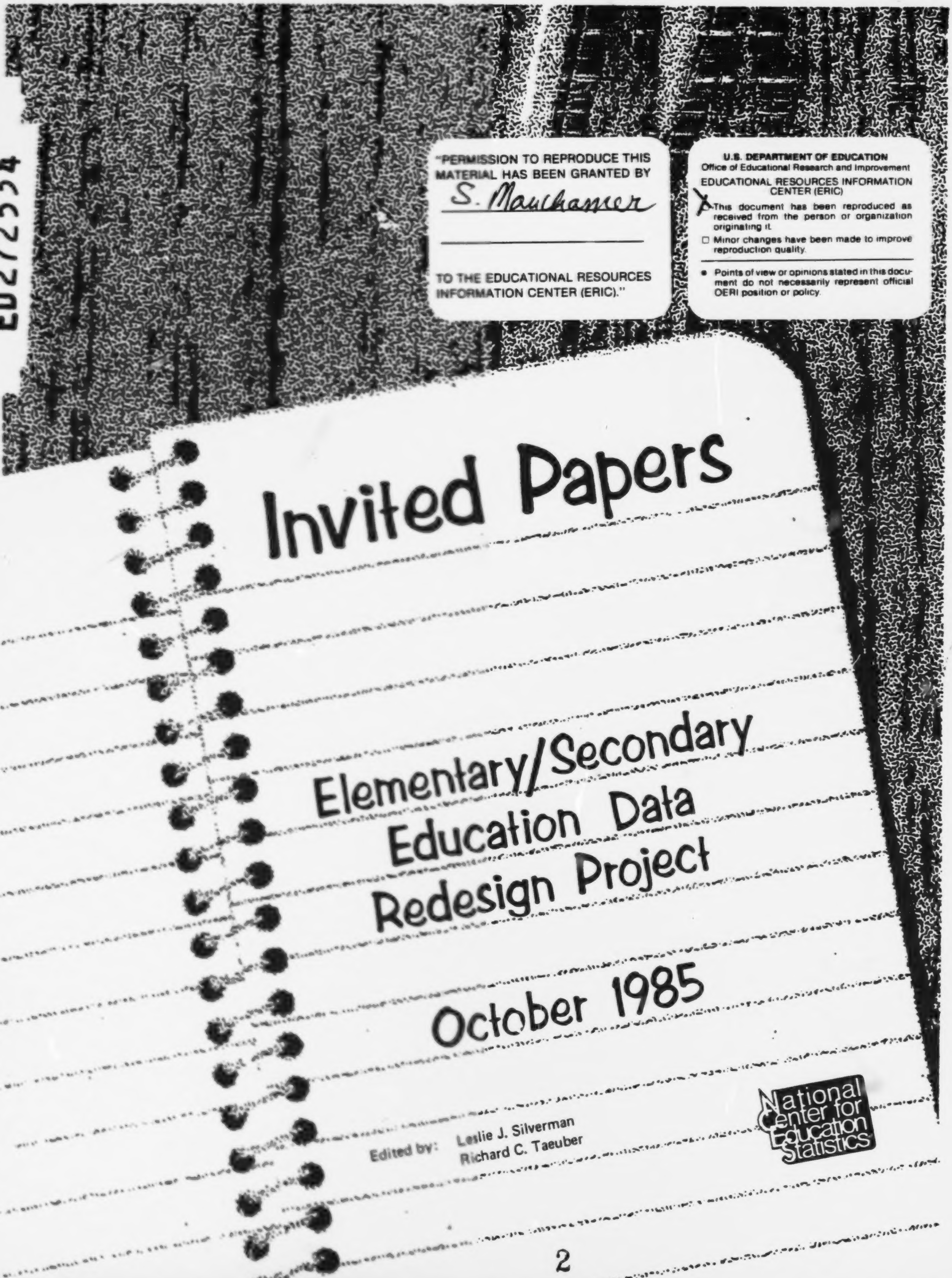
****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED272534

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
S. Mauchamer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☒ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.



Invited Papers

Elementary/Secondary Education Data Redesign Project

October 1985

Edited by: Leslie J. Silverman
Richard C. Taeuber

National Center for Education Statistics

TM 860 450

PREFACE

In early 1985, the Center for Statistics began a thorough review of its data collection program as the Elementary/Secondary Education Data REDESIGN Project. The purpose of this review is to address questions of the suitability, scope, reliability, and timeliness of CS' statistical data in terms of the purposes those data are expected to serve -- purposes relative to instructional and administrative needs, and to education policy issues. The result of this effort may well be the design of a new plan for national data collections from institutions and individuals to be implemented over several years.

To initiate the dialogue, a number of individuals and organizations where invited to prepare a paper following guidelines (over leaf). The invited individual authors were persons likely to (a) use quantitative approaches in their analyses, and thus be data users, and (b) have the interest and time to produce a "thought paper", not a research paper. There was no intent to represent known positions or to provide balanced representation across sectors of the education community.

The organizations were selected because of an identification with education issues and concerns, and were invited to participate in any or all stages of the REDESIGN Project, including the option of writing a paper.

38 invited individual papers, 18 organization papers, and 3 agency papers are included in this volume.

To facilitate public comment in response to, or as stimulated by the initial set of papers, CS asked a team of writers to distill the essence of the papers. The product of that effort was a report: Synthesis of Invited Papers: A Public Discussion Draft (NCES 85-114). The purpose of the Synthesis report was to represent in concise, coherent manner the breath and content of the invited papers which had been received by August 10, 1985. In that way, participants in the REDESIGN Process were not required to read the full set of invited papers.

The Editors

Guidelines provided authors and other participants in the REDESIGN Project

QUESTIONS BEING ADDRESSED

The authors were provided a guidance set of questions, but advised that those questions were not meant to be restrictive, rather guide lines for the types of comments of interest to the redesign project. Further, they were asked for possible survey questionnaire items, specific measures, or indicators that could improve future analyses, and changes to existing procedures, definitions, and coverages.

The guiding questions were:

1. *What data or data series are needed to support deliberations on future policy issues, or decisions on instructional and administrative needs, during the remainder of the 20th century? (Link the issues or needs to the data items.)*

2. *What additional data or data modifications – in items, measures, indicators, or sample universes/frames – would improve the utility, validity, or reliability of current national data files? (Identify the data files and how they would be improved.)*

3. *What current NCES data series are most important to maintain and why?*

4. *What current data elements or series are recommended to be deleted from current data programs and why?*

5. *What other suggestions are offered for improving the relevance, technical quality, and utility of the NCES data programs?*

ABOUT THE PAPERS

The charge given to all authors and organizations preparing papers was very general in nature. All were encouraged to go beyond any one specific issue or area of major personal concern. Authors have, in fact, been encouraged to represent the breadth of issues in elementary and secondary education.

> For this review – "Elementary and Secondary Education" can include educational experiences from birth through the transition to postsecondary education or the workplace. The review can include all educational experiences: public or private, in "schools" or other locations, organized or not.
>
> For this review – "National Data" can include reference to any data collections, existing or proposed, and need not be confined to NCES data programs. For example, the paper can discuss any Federal Government or other national education data programs, such as the National Assessment of Educational Progress (NAEP), the Survey of Local Government Finances (Bureau of the Census F-33 Series), or the Decennial Census of Population and Housing.

BEST COPY AVAILABLE

INDEX

PAGE	AUTHOR	TITLE
1	James M. Banner, Jr.	"Revising Educational Statistics"
6	Stephen M. Barro	"NCES Data on School Finance and Teachers: Assessment and Recommendations"
40	Sue E. Berryman	"Education and Employment: Substitution Possibilities and the Teacher Labor Force: Supply and Demand"
58	John H. Bishop	"Data Collection for Improving Elementary/Secondary Education"
89	Alphonse Buccino	"Monitoring the Condition of Education"
100	James S. Coleman	"Data Needs for School Policy in the Next Decade"
113	Joseph M. Cronin	"Issues in National Educational Data Collection"
126	Linda Darling-Hammond	"Data on Teachers and Teaching: Opening the Block Boxes of Education"
142	Jane L. David	"Improving the Quality and Utility of NCES Data"
149	Eugene E. Eubanks	"Data Needs for Big City Schools"
152	W. Vance Grant	"An Elementary and Secondary School Statistics Program for the National Center for Education Statistics"
160	Jane Hannaway	"Two Suggestions for NCES Data Collection"
174	Carole Hall Hardeman	"The Quest for Excellence/Pupil Self-Esteem"
191	Forrest W. Harrison	"Review of Elementary/Secondary School Data: Needs of the National Center for Education Statistics"
200	Willis D. Hawley	"Educational Statistics and School Improvement"

INDEX (cont. 2)

PAGE	AUTHOR	TITLE
232	Richard H. Hersh	"Organizational Efficacy as a Research Focus for School Improvement"
247	Paul T. Hill	"The Politics of Educational Data Collection"
257	Asa G. Hillard III	"Information for Excellence and Equity in Education"
276	Maureen McClure and David N. Plank	"Educational Statistics for Educational Policy: A Political Economy Perspective"
293	James M. McPartland, Henry Jay Becker, and Robert L. Cain	"A Model for N.C.E.S. Research on School Organization and Classroom Practices"
331	Inabeth Miller	"A House of Bricks"
340	Richard J. Murnane	"Priorities for Federal Education Statistics"
365	Gary Natriello	"Products and Processes of the National Center for Education Statistics: An Agenda for the Next Decade"
379	Allan Odden	"Federal Collection of School Finance Data: New Needs for an Era of Education Reform"
392	Penelope L. Peterson	"The Elementary/Secondary Redesign Project: Assessing the Condition of Education in the Next Decade"
407	Valena White Plisko, Alan Ginsburg, and Stephen Chaikind	"Education Statistics: Assessing National Data"
431	Elizabeth R. Reisner	"New Areas for Educational Data Collection: What Students are Taught and What They Learn"
445	Susan J. Rosenholtz	"Needed Resolves for Educational Research"
470	Diane Scott-Jones	"Assessing American Education: Shrinking Resources, Growing Demands"
495	Ramsay W. Selden	"Educational Indicators: What We Need to Know That We Don't Know"

PAGE	AUTHOR	TITLE
513	Marshall S. Smith	"Thoughts on Improving the Quality and Utility of NCES Data"
519	Gail E. Thomas	"Issues and Considerations for a Ten-Year Program on Elementary and Secondary School Data Collection"
531	Brenda J. Turnbull	"Comments on an Elementary and Secondary Education Data Program"
541	Michael D. Usdan	"Educational Data Needs for the Balance of the 20th Century: Some Perspectives on the Emerging Environmental Context"
565	Rafael Valdivieso	"Hispanics and Education Data"
574	Herbert J. Walberg	"National Statistics for Improving Educational Productivity"
599	Margaret C. Wang	"Toward Achieving Educational Excellence for All Students"
625	Louise Cherry Wilkinson	"Assessing Students' Social and Communicative Achievements in School"

ORGANIZATIONS

PAGE	AUTHOR	
633	American Association for Counseling and Development	'A letter'
639	American Association of College for Teacher Education	'Written comments'
653	American Library Association	'A letter'
655	American Statistical Association, Robert G. Lehnen	"Educational Statistics for Studies of Policy and Administration"
664	Association of American Publishers	'A letter'
667	Association of Teachers Educators, Lee Bartolini	"The Identification of Data Needed to Support Deliberations on Policy Issues"

INDEX (cont. 4)

ORGANIZATIONS

PAGE	AUTHOR	
672	Council of Chief State School Officers	'Written comments'
680	Council of Chief State School Officers	'A letter, July'
682	Council of Chief State School Officers	'A Letter, September'
684	The Council of State Governments, E. Norman Sims and Deborah A. Gona	"Assessing the Education Statistics Information Needs of Non-SEA Public Policy Decision Makers"
699	Educational Testing Service, William W. Turnbull	"Needs for Data in Education"
719	Lutheran Church-Missouri Synod	'A letter'
721	National Center for Research in Vocational Education	"Data on Vocational Education: Problems and Recommendations"
730	National Education Association	"Suggestions for the NCES Redesign Project"
748	National Governors' Association	"Comments on the NCES Redesign"
762	National School Boards Association	'A letter'
766	Rural Education Association, Bruce Barker	"Research and Data Needs for Small/Rural Schools"
774	State Higher Education Executive Officers	'A letter'

AGENCIES

PAGE	AUTHOR	
778	Equal Employment Opportunity Commission	'A letter'
780	National Science Foundation	'A letter'
783	Office of the Assistant Secretary of Defense (Force Management & Personnel)	'A letter'

REVISING EDUCATIONAL STATISTICS

James M. Banner, Jr.

Senior Research Associate

Council for Basic Education

It is easy enough, as many have shown, to find serious deficiencies in the gathering and presentation of educational statistics. Yet in reflecting on flaws in existing data and on the consequences of their weakness, we often lose sight of the general principles that should guide any improvements and revisions of them. As a prelude to some specific concerns and at the risk of reviewing some hackneyed themes, I should like to consider these principles, of which in my view there are five.

The first is that of accuracy. The raw data used to construct statistical series must be valid and complete; their weaknesses and limitations must be specified; and probable margins of error and levels of confidence must be stated. Only in this way can faith in the strength of the statistics be raised, their utility increased, and ill-conceived employment of them -- for ideological, propaganda, and other ends -- be curbed.

Second to accuracy in order but not in importance is the principle of comparability. Accurate data to not assure their utility unless comparable information has been sought in comparable ways in comparable forms from comparable sources. The failure to create comparable information subverts all efforts accurately to assess differences and changes in educational circumstances between places and over time and thereby weakens or nullifies policies built upon them.

The third principle is that of potential utility. Educational data should be collected even in the absence of contemporary or immediate needs in policy-making, and they should be collected in such form that they can be reformulated for changing uses. Failure to collect data that may seem superfluous for immediate presentation can reduce the effectiveness of future policies by inhibiting their substantiation and the evaluation of their historical significance.

Related to potential utility is a fourth principle, that of history. Too often ignored but essential to the evaluation of the meaning of any changes over time, historical statistics provide the context for the interpretation of long-term change. Without historical statistics, all claims about trends, the significance of data, and the solutions to problems are suspect.

Fifth and final is the modest, but essential, principle of ease. If responses to requests for data become burdensome, or if data, once collected, cannot be easily used or assessed, statistics decline in appeal. Therefore, data reporting must be made simple for those from whom it is solicited, and data samples should be collected in place of statistics on a complete universe whenever possible. Relatedly, data that lend themselves readily to scholarly analysis -- even to such basic and simple uses as correlations -- should be collected so that the function of NCES can remain primarily that of data collection rather than analysis, which should be carried out by others. Furthermore, data should be presented in such a way that local reporting units, be they individual schools or entire school districts, can assess their relative positions in comparison with averages for similar reporting units.

Taken together, these principles, if applied, will help ensure the integrity and applicability of the statistical information upon which we must in part base all general educational policies. Yet they will not at all times be of equal moment. Occasionally, circumstances may require emphasis upon a single one in order to increase confidence in the strength of all data. In my view, now is such a time. What requires immediate attention, in order to secure the utility of what information, however accurate, we now have, is the comparability of educational data. And the Department of Education must take the lead in this effort.

For many reasons, the American public now seeks to be assured of improvements in education at all levels, especially in the primary and secondary schools. And, characteristically, it wants information that compares present conditions with those of the recent past and conditions in one jurisdiction with those in others. Yet the plain fact of the matter is that the data available to provide such comparisons are embarrassingly weak. The public is being misled by their use. And policies based upon them may therefore be without foundation or desired effect. For these reasons, even before attempts are made to improve the collection of accurate raw data, the data that are solicited must be collected in ways that allow accurate comparisons of educational conditions among jurisdictions and across time. This is not to say that the search for improved accuracy should not proceed. Rather, the public as well as policy-makers must be assured that, even in the absence of reliable data, the nature, direction, and extent of change in American education is accurately represented by the data that we do have. That cannot be done now.

Moreover, the data that are gathered and published must be consistent over time. Too often, the existing data series are presented differently, due both to altered data-gathering methods and changed survey questions, from one year to the next. This renders virtually nugatory all attempts to evaluate changes in educational conditions.

How should the endeavor to improve the comparability of data be undertaken? Despite the sensitivity, as misplaced as it often is, toward the centralization of educational policy, the attempt to collect comparable data should be led by a national agency or organization, if not by the Department of Education then under the auspices of the Council of Chief State School Officers. In any case, both must act in concert with state officials to standardize data collection and presentation.

One problem facing the Department and NCES is the use of proxies for collecting data, as a result of which the government's need for information depends upon the needs of other organizations and agencies and often falls victim to them. This is most visible in data from sample surveys, often undertaken at a third remove -- that is, at the behest of clients of reporting organizations. If adequate and consistent proxy surveys and data cannot be secured, then the Department must try to secure them directly or by contract.

Additions to, and changes in, existing data series will not be of much use in my estimation without speedy attention to these general and encompassing obstacles that now stand in the way of sound data collection. Nevertheless, some more specific matters also require tending to if the Department's data series are to be made more useful. Omissions in the data, especially those concerning the quality of primary and secondary schooling, must be filled. Although all efforts to measure the quality of education must end in approximations, they must nonetheless be made and the glaring lacunae in existing data closed.

For instance, despite the intensive debate regarding the preparation and qualifications of school teachers, we have no adequate current or long-term information to that effect. Data series, by state, about teachers' levels of education (bachelor, master, and doctoral degrees and credit hours beyond certification), permanent and temporary certification, the ratio of those teaching with and without certification in their field of instruction, the percentage of those who have switched teaching fields, the percentage of those teaching subjects in which they were not initially trained, student-teacher ratios by fields, the results of newly-instituted competency tests, and similar matters should be maintained and presented historically. Similarly, we should have measures of scores on the College Board and other advanced placement and achievement tests, broken out by states and public and private schools. Efforts to measure changes in college entrance requirements should be endeavored too, as well as changes in the relative proportion of advanced and AP courses and general courses and in the number and proportion of students taking each.

In many cases, data are unwisely aggregated or aggregated in forms that reduce their usefulness. Despite jurisdictional realities, a better unit for comparison of much data is probably the standard metropolitan area rather than the state, at least for urban schools. Analogously, data regarding public and private schools should be distinguished. Much of the data fail to reflect the dual system of American education at all levels; and even those few efforts to distinguish between different kinds of private and religious schools are not carried out consistently in the series now published.

Discontinuities in students' school experiences are not adequately assessed either. Measures of dropping out, notoriously weak as they may be, fail to distinguish between chosen interruptions in schooling ("stopping out") and involuntary interruptions (such as moving from one school to another). These data could be collected via sample surveys and with the assistance of the Bureau of the Census. In fact, one measure of the stability of schools might be a simple ratio of new to returning students per jurisdiction.

Of the curricular information that should be added to existing data series, that concerning the state of art and music should have high priority. The omission of these subjects from consideration by authors of the recent flood of national reports on schooling was egregious. By contrast, already collected data on so-called remedial courses are probably the weakest that are published, given the wide differences in definitions of remediation by school districts and individual, especially private, schools. Such data could be omitted.

Finally, some attention should be given to publishing revised data as better information is collected or errors in previously published data are discovered -- as, for example, the Treasury Department routinely does with its periodic economic statistics. For example, it is exceedingly unlikely that the percentage of 18-19 year olds graduating from high school rose by 5.4 per cent between 1975 and 1976 and then dropped 3.3 per cent the next year. Chances are that the data for 1976 are suspect. Such evident abnormalities should be carefully checked; and if the data are found to be erroneous, they should be revised in all subsequent publications.

In all of these modest ways, and in others that will be proposed, the foundations of knowledge about American education will be greatly strengthened.

SMB ECONOMIC RESEARCH, INC.
6315 29th Place, N.W.
Washington, D.C. 20015
(202) 362-0316

NCES DATA ON SCHOOL FINANCE AND TEACHERS: ASSESSMENT AND RECOMMENDATIONS

Stephen M. Barro

revised
August 1985

This paper has been prepared for the National Center for Education Statistics under purchase order NIE-P-85-9521. The views expressed are solely the author's and do not necessarily reflect the opinions or policies of NCES.

INTRODUCTION

In this paper, I comment on the two categories of NCES elementary-secondary data with which I am most familiar from my own work: data on school finance and data on teachers. In discussing each, I review current NCES offerings and their limitations, identify additional data needed to serve policymaking and research, and suggest modifications (mainly expansions) of NCES data collection efforts.

The two data categories I deal with, school finance and teachers, may seem at first to be only distantly related to one another, but the connection between the two becomes more apparent when one thinks of teacher data as a subset of data on educational resources. There is a logical chain of variables running from money to resources to services to educational results, and the two categories considered here correspond to adjacent links in that sequence. Information on school finance indicates how money for education is raised and how it is expended; information on resources (specifically, teachers) indicates what money buys and how the inputs into schooling are used. As I will argue below, the failure of current data to illuminate the finance-resource connection is a major shortcoming of the existing NCES data system.[1]

The following comments are limited in several important respects. First, I have stopped well short of proposing specific data system designs, both because that task is too large for this type of paper and because of my own limited expertise. Second, I have not addressed technical questions of data quality because I have nothing to add on that subject to what NCES already knows. Third, I have not taken budget constraints into account in discussing what data NCES should collect. If that gives my suggestions an air of unreality, I apologize, but I thought that someone should consider, without prior restraints, what information an advanced country might reasonably want to gather about its own educational system.

---

[1] I would have preferred to look at resource data across the board but did not have time to undertake that task. However, much of what is said here about teachers applies to instructional personnel generally and hence to the most important resource category.

SCHOOL FINANCE DATA

NCES has long been in the business of collecting school finance data and reporting basic finance statistics for states and, at times, for local school districts. These data have often been used for making gross fiscal comparisons (e.g., of per pupil spending among states) and charting broad trends in support for the schools, but they have been of little use for addressing the central "adequacy" and "equity" issues of school finance, analyzing resource allocation patterns, or relating funds and resources to educational results. My main purpose in this section is to explain why the existing finance data are inadequate for such purposes and to suggest what it will take to make them more useful.

Current NCES Finance Data

The NCES currently produces what might fairly be described as skeletal information on school finance. The principal data collection instruments, the Common Core of Data (CCD) State Fiscal Report and Public School District Finance Report, distinguish among revenues from local, intermediate, state, and federal sources and among outlays for instruction, support services, and noninstructional services. They also break out, on the revenue side, receipts from property taxes, tuitions, and intergovernmental transfers and, on the outlay side, spending for salaries, employee benefits, debt service, and construction. State-by-state data are reported in the Condition of Education and Digest of Education Statistics (hereafter, "Condition" and "Digest," respectively), typically with lags of three to four years.[2] Financial data for selected large local districts have been published irregularly in the past, but the latest such data to appear in the Digest are for the 1979-80 school year.[3] Although district-level data have supposedly been collected annually in the CCD surveys, neither the data themselves nor any findings based upon them (except for state totals) have, to my knowledge, been published. I do not know whether or for what purpose such data are used once they have been collected. In particular, I note that there are no NCES publications describing the distributions of revenues or expenditures among local school districts, either nationally or within states, even though distributional statistics (especially indicators of intrastate disparities) have long been the central concerns of school finance policymakers and researchers.

It is also notable that the categories used by NCES to collect and report state and local finance data are now less detailed than in earlier years. Formerly, distinctions were made on the expenditure side among such traditional school accounting categories as instruction, administration, plant operation and maintenance, fixed charges, etc. Now there appears to be only the three-way CCD breakout indicated above. This is only a small loss, however, since as I will discuss below, the expenditure information of greatest interest is, and always has been, concealed within the overbroad "instruction" category.

---

[2] For example, the Digest, 1983-84, presents revenues and expenditures for 1980-81 (pp. 78-79), and the Condition, 1984, gives expenditures and revenues for 1981-82 (the latter, for some reason, from NEA estimates rather than NCES' own sources).

[3] Digest, 1983-84, pp. 60-61. The identical 1979-80 finance data also appeared in the 1982-83 edition.

To put the NCES efforts in context, one must also take into account the separate system of school finance data collection and reporting operated by the Governments Division of the Census Bureau. The annual Census survey (a census in some states and a sample survey in others) covers most districts in the U.S. The Bureau's reports, entitled Finances of Public School Systems, present expenditure and revenue data for states and for individual school districts of over 15,000 enrollment. They are produced with a time lag of under two years (i.e., the 1982-83 data have appeared). The categories used, while generally no more detailed than those of NCES, reflect certain important distinctions not made in the NCES reports (see below).

Data Shortcomings and their Consequences

There has been much discussion over the years of certain technical problems concerning the school finance data, including problems of coverage, definition, and data comparability (e.g., in the treatment of pensions). These issues are familiar to NCES staff, and there is little I could add that would be helpful. I focus instead on what I believe are some broader issues bearing on the usefulness of the finance data for research and policymaking. I consider, first, the lack of sufficient expenditure detail to make the connection between finances and resources; second, the absence of distributional statistics pertinent to major school finance concerns; third, gaps in the information on school revenue; and finally, the problem of instability over time in school finance data collection.

The Lack of Expenditure Detail and the Disjunction between Expenditure and Resource Information

The education expenditure data currently reported by NCES are serviceable, at best, for making gross fiscal comparisons among states and examining broad trends in public support for the schools. Even in those applications they can be misleading, because differences in dollar outlays among states and over time do not necessarily correspond to differences in educational resources. But more important, they are not suitable for other purposes to which potential users would like to apply a school finance (or finance/-resource) data base. Setting aside the distributional issues for separate consideration, these applications include such things as the following:

- Analyses of what education money is used for (i.e., what money buys) in different states and LEAs,
- Comparisons of amounts spent and prices paid for particular kinds of resources (especially teachers and other instructional staff) in different places or at different times,
- Analyses of amounts expended for different levels or types of instruction, or on behalf of different categories of pupils,
- Research on relationships between school spending and educational services and outcomes.

Two reasons for the limited usefulness of current data are that expenditure data are not collected in sufficient detail to be connected with resource categories, and expenditure and resource categories are not coordinated. Consequently, information on dollar outlays cannot be linked to anything real. Most expenditures of direct educational interest, in fact, are contained within the single, overbroad, traditional category, "instruction," which is not decomposed either by type of resource or by the various purposes for which instructional resources are used. Other costs of instructional resources, notably fringe benefits, are hidden within the mysterious category, "fixed charges." In consequence, one cannot tell how much of a state's or an LEA's education budget is expended on classroom teachers, as opposed, say, to administrators, specialists, or instructional materials; nor how much is spent for staff compensation, counting nonsalary as well as salary costs; nor what is spent in high schools, as compared with elementary schools, or for vocational, as opposed to academic, programs; nor, in comparing states, whether differences in instructional outlays per pupil are due to differences in staffing ratios, differences in salary per staff member, or both. Without being able to make such distinctions, one has very limited ability to make sense of interjurisdictional financial comparisons or to address the issues of resource and financial adequacy.

There is little doubt about what is needed to create these missing analytical capabilities. In general, the expenditure data need to be disaggregated and rearranged into educationally and economically meaningful categories. For instance, to connect finances to resource allocation (and, ultimately, to services and effectiveness), it would be necessary to disaggregate expenditures (especially but not exclusively outlays for "instruction") into appropriate resource, or "object," categories. Specifically, I envision a system of combined expenditure and resource accounts, in which instructional outlays are explicitly linked to numbers of instructional staff, amounts of other resources, staff compensation, and other prices in a manner something like the following:[4]

Instruction

Instructional personnel

Classroom teachers: (number) x (avg. salary) = (outlay)

Specialist teachers: (number) x (avg. salary) = (outlay)

Teaching aides: (number) x (avg. salary) = (outlay)

... etc ...

Instructional materials: (outlay)

Total Instruction: (outlay)

Given such a data set (especially with the data expressed in per pupil form), it would be immediately apparent whether differences in spending levels among places or over time reflect differences in real resources or differences in salaries and other prices. If nothing else, this should promote more meaningful interstate and intertemporal comparisons. But going beyond mere comparisons, I would expect such data to make possible a variety of analyses of resource allocation patterns and behavior now feasible only with ad hoc data bases from selected states. One would be able to ask, for example, what shares of the incremental education dollar tend to go for more teachers, higher-paid teachers, administrators, etc. and how those shares vary among school districts with different characteristics. Thus, expansion of the expenditure data along the lines indicated should lead to an improved understanding of what money buys and how the inputs into schooling are likely to be affected by changes in finance.

---

[4] For simplicity, I treat personnel cost in this example as synonymous with salary, whereas more appropriately it should be measured by total compensation, which in turn should be decomposed into current salary, deferred compensation (retirement programs) and other fringe benefits.

A very different question from what money is spent on is what it is spent for--i.e., for what purposes or programs or on whose behalf. There has long been interest, for example, in how funds and resources are distributed between the elementary, middle, and high school levels, among programs (e.g., academic versus vocational), and among "target groups" (handicapped, disadvantaged, limited-English proficient), but the data have not been available to provide answers. In this case, however, it is not so easy to fault NCES for failing to provide relevant breakdowns of outlays because, except for the relatively straightforward breakdown by level of school, it is not clear that valid disaggregation is feasible.

The problem, basically, is that school districts do not maintain the kinds of cost accounting systems that would be needed to attribute costs to particular pupil categories or programs. Segregating such costs would be easy if, for example, disadvantaged and handicapped pupils were educated separately from regular pupils and vocational pupils separately from academic pupils, but such is rarely the case. Typically, the different types of pupils attend the same schools and many of the same classes, are served jointly by at least some of the same instructional staff, and share the same facilities and support services. Under such conditions, it would take very detailed time and cost accounting systems to allocate outlays properly.

Given that few districts operate such accounting systems, an attempt by NCES to collect outlay data by pupil category or program would probably do more harm than good. Reporting by local and state officials would probably be based more on the availability of earmarked state or federal funds for particular programs or groups than on actual costs incurred. The resulting data would be unlikely to represent properly either the total costs or the "add-on" costs of the programs in question. I conclude, therefore, that whereas disaggregation by level of school is feasible (because costs are incurred by distinct organizational units), disaggregation by program or pupil category is impractical at this time. Acquiring the ability to disaggregate by the latter categories should be treated as a longer term goal, to which NCES methodology development resources might reasonably be devoted. Meanwhile, disaggregation by level is worthwhile in its own right (e.g., so that data on high school resources—not high school and elementary resource combined—can be juxtaposed to data on high school outcomes) and should be pursued as a shorter-term goal.

The Lack of Statistics on Distributions of Spending and Revenue within States

Unlike aggregative financial data for states, financial data for local school districts are not meaningful or accessible until they have been analyzed and summarized statistically. Users cannot scan local expenditure and revenue data for thousands of districts in search of patterns, as they can the data for 50 states. Nor can they rearrange, manipulate, and summarize the district-level data to suit their needs, unless they are equipped (and funded) for large-scale data analyses. The value of the district-level finance data to most users depends, therefore, on how extensively and appropriately such data are analyzed and summarized statistically by NCES itself.

Different kinds of summaries and analyses are useful, of course, for different purposes. NCES, in the past, has provided certain summary statistics of district-level finances (notably, in publications entitled Statistics of Local Public School Systems, Finances). For instance, it has compared levels of spending among size strata of districts, among geographical regions, and among central city, suburban, and nonmetropolitan districts. Such analyses have not recently appeared, however. They are moderately useful, and it would be helpful if they were revived or, preferably, expanded to cover additional district characteristics and cross-classifications.

Far more important, however, is that NCES has always studiously avoided producing the statistics wanted most by scholars and policymakers involved in school finance, namely, statistics on the distributions of expenditures and revenues among school districts within states. Such statistics (e.g., indicators of intradistrict inequality in school spending and tax rates), though well within NCES' technical capacity to prepare, have been considered too politically sensitive by NCES leadership. Their publication might offend some state education agencies (presumably those of high-disparity states), and they could even be used as evidence in Serrano-type school finance lawsuits. Thus, NCES has feared that developing such statistics, or even facilitating their development by others (e.g., by making available edited, "user-friendly" district-level data files), could undercut the state-agency cooperation on which NCES depends for much of its access to state-local data.

The NCES position regarding distributional statistics was brought out most clearly in the well-known (among school finance specialists)

"Profiles" episode of 1979-80. At that time, federal interest in school finance equity had reached a peak, and NCES had been directed explicitly by the Congress (P.L. 95-561, Sec. 1201) to produce a set of "profiles of state school finance equalization," showing the extent of fiscal disparities and deviations from fiscal neutrality in each state. This profiles report was to be the first of an ongoing biennial series. NCES did not want to produce such a document, primarily, as I understand it, for the reason alluded to above. Its avoidance strategy was one of protracted technical delay. Numerous outlines and drafts of a profiles report were prepared, some by contractors and consultants and some by NCES staff, all to be sent back for reworking or rejected outright by internal review committees. The strategy worked. By late 1980, the federal interest in school finance equity had declined precipitously, and NCES was able to fulfill its formal obligation by sending a report to Congress, while suppressing its further publication. Needless to say, nothing has been heard of since of the permanent biennial series of profiles reports; nor has anything else emerged from NCES (despite the availability of the CCD district-level data) on patterns of school financing within the states.

This history raises several issues. First, of course, there is the specific question of whether NCES should abstain from producing statistics pertinent to the central concerns of school finance (or, for that matter, to any other educational concern) because such publication might displease some states. To gain some perspective on the issue, consider what the reaction would be if the appropriate federal agencies failed to produce data on infant mortality, crime rates, inadequate housing, poverty, or for that matter, the educational attainment of the population because officials in the low-ranking states might be upset. It is unlikely, in such areas, that state embarrassment, displeasure, or even the implicit threat of future noncooperation would be deemed acceptable reasons for delimiting the federal government's information gathering role. The same, in my view, should be true in education and, in particular, in school finance.

There is, second, the broader issue of NCES' posture vis-a-vis the states. Often, NCES has gone to state agencies as a supplicant, seeking approval to collect even data that (a) bear directly on matters of national concern or (b) are needed to produce a coherent picture of the nation's education system. Although I cannot prove it, I believe that the sparsity of NCES data on finance, teachers, and other aspects of education is due in considerable part to this excessive diffidence. Whatever the merits, or imperatives, of a supplicant posture may have been in the past, there is now good reason to reconsider. A conservative administration has chosen to disseminate, for the first time, state-by-state comparative data on pupil performance, despite the unhappiness this doubtless causes states at the bottom of the performance ladder. In the area of finance and elsewhere, NCES might well profit by that example.

Third, there is an issue of less social significance but considerable importance to the future role of NCES, namely, how far the agency should generally be expected to go in analyzing, as opposed to collecting and compiling, education data. Current NCES practice in that regard is very uneven. For instance, the agency has sponsored many analyses of its longitudinal survey data (the National Longitudinal Survey and High School and Beyond), even extending to the development of sophisticated behavioral models, but has not supported even rudimentary analyses of the school finance data. I cannot offer a general recommendation regarding the appropriate limits on the analytical role. The answer obviously depends, among other things, on the agency's mission in a reorganized Office of Educational Research and Improvement. It seems to me, however, that there is one valid threshold criterion: NCES should be expected to undertake enough descriptive (as opposed to inferential) statistical analysis to present meaningful syntheses and analyses of patterns in its major publications. This has not been done in either of the two areas covered by this paper.

Deficiencies in NCES Revenue Data

Published NCES data on public school revenue are useful mainly for making interstate comparisons and examining broad trends in revenue levels and shares of revenue from federal, state, and local sources. They are of little use for any further analysis of revenue sources or revenue-raising instruments. I suggest here three steps that would increase substantially the usefulness of the revenue information: elaboration of the revenue categories, production of data on revenue bases, and compilation of nonstatistical, descriptive data on state-local school revenue systems.

Revenue Categories. The three-way classification of revenue by federal, state, and local sources is useful as far as it goes, but it conceals some important, policy-relevant information about the forms and channels through with revenue is obtained. The following changes in revenue categories would provide a more accurate and complete picture. First, revenue from local sources should be decomposed into tax revenue (distinguishing between property taxes and other taxes), current charges or user fees (mainly tuitions), contributions from parent governments, and "other revenue."[5] The distinction between taxes raised by independent local districts and the contributions, or appropriations, that dependent districts receive from their parent counties, municipalities, or townships is especially important (and has long been recognized in the Census Bureau's Finances of Public School Systems). Second, research on school finance systems and school finance equity would be aided if distinctions were made between general, or unrestricted, state aid to local districts and categorical, or restricted, aid (the latter including aid earmarked for particular programs, beneficiaries, or objects of expenditure). It would also be helpful to arrange the data on state and federal aid in a manner that brings together both federal and state support for closely related special programs. Third, a distinction should be made between direct federal education aid to local agencies and aid that is "passed through," and often redistributed by, the states (this distinction is standard in the Census publications).

Revenue Bases and Tax Rates. The relationships among revenues, revenue bases, and rates of taxation for schools (effort) are important questions in school finance, but NCES does not provide data on bases or rates. Thus, one cannot use NCES data to determine how variations in spending or tax rates among districts are influenced by variations in local tax bases or, from the equity standpoint, to determine the degree to which school spending is a function of local wealth. During the late

---

[5] Some of the items mentioned, including receipts from property taxes and tuitions, are already included as "special exhibits" in the CCD surveys (according to NCES, "The Elementary/Secondary Education Data Acquisition Program," an unpublished summary description of NCES elementary-secondary data acquisition systems).

1970's, NCES did attempt to produce data on equalized property tax bases by district, but the effort was beset by technical difficulties and has not been repeated. The technical difficulties are indeed substantial, since each state has its own methods of defining and measuring the tax base, and it would be undesirable to publish a national set of tax base data before the main problems of comparability are overcome. In the interim, however, NCES could provide a useful service by reporting relative local tax bases and tax rates, using the official definitions of each state. Although this would not allow interstate comparisons of wealth, it would support analyses of intrastate relationships between spending and wealth and comparisons of the findings from such analyses among states. Thus, the ability to analyze school finance systems could be enhanced significantly at very little cost to NCES.

In the longer run, NCES could make a valuable contribution to school finance research by producing not only comparable data on local wealth but also data on a broader range of local economic variables. The mapping of decennial Census data by school district is important in this regard because it provides information on income and poverty in each district and on many related demographic characteristics. Selected Census data items should be merged with NCES' school finance data and made available as a user file. With such a data set, it would be possible to link fiscal and resource variables to multiple district attributes and to address a variety of concerns about how different communities finance their schools.

Descriptive Information on Revenue Systems. Although NCES is a statistical agency, it is well within its charter to collect and compile the nonstatistical, descriptive information needed to make sense of its numerical data. In the area of school revenue, two bodies of such descriptive information are important: descriptions of state aid mechanisms and formulas and descriptions of each state's legal framework governing the raising of local school revenue. Neither type of information is now provided by NCES.

Over the years, some information on these matters has been produced by other offices of the Education Department. Specifically, a series of volumes edited by Esther Tron and published irregularly by the Department provided state-by-state descriptions of school finance systems, and materials published by the Education Commission of the States (ECS), under sponsorship of NIE, provided tabular summaries of system characteristics.[6] For various reasons, however, neither effort has filled the requirement for comprehensive, consistent, and timely descriptions of how each state raises its public education funds.

It seems reasonable that NCES, as the information-gathering arm of the Department, should take on the responsibility for maintaining a continually updated file of such information. Specifically, I suggest that NCES should determine and publish on a regular annual basis the method used by each state to fund its schools (including full mathematical details of the formulas) and the rules under which each state's LEAs are permitted to raise revenue (i.e., taxing authority, definitions of tax bases, fiscal constraints, referendum requirements, etc.). With that information in hand, analysts and policymakers would be much better situated to interpret and use the improved and expanded financial statistics called for above.

---

[6] The most recent of the Tron compendia is Public School Finance Programs, 1978-79, U.S. Government Printing Office, Washington, D.C., 1980. The ECS tabulations are published as a series of wall charts, entitled School Finance at a Glance, the most recent of which is for 1983-84.

Stability Over Time

Apart from any deficiencies of content, an NCES practice that has detracted from the value of the school finance data is allowing the data production effort to fluctuate from year to year in response to political currents. Most recently, when federal interest in issues of school finance adequacy and equity plunged toward zero in 1980-81, NCES cut back its school finance effort, reducing the coverage and detail of its surveys as well as its data processing and publication activity.

This is short-sighted behavior. Although the federal demand for school finance studies is now nil (there does remain, however, substantial state-level interest in school finance adequacy, equity, and reform), experience indicates that interest in such topics is cyclical. The surge of interest that led to the aforementioned Profiles requirement and a Congressionally mandated school finance study in 1979-80 was preceded by another such surge in 1971 (the President's Commission on School Finance). It is safe to say that interest in the topic will rise again in the future. When it does, NCES will find itself with (a) large holes in the historical data base, due to failure to sustain its basic surveys, (b) antiquated data systems, due to lack of development effort, and (c) no capability to generate new data without a substantial time lag.

This on-again, off-again behavior makes little sense with respect to an area as fundamental to policymaking as school finance. The continuity of the data base is as important a determinant of its usefulness as is data quality. I recommend that stable, annual collection and publication of financial data (preferably with a much-reduced time lag) be adopted as an agency norm under the new ten-year data improvement plan.

DATA ON TEACHERS

The recent upsurge of interest in teacher quality, teacher compensation, and teacher supply and demand has drawn attention to an area in which NCES' efforts are particularly weak: the production of data on teachers and teaching. Current NCES data on these subjects do not suffice even to provide general background information pertinent to policy concerns, much less to support research on problems and possible solutions. Moreover, although demands for information on teachers are now unusually intense, they cannot be characterized as unanticipated or "new." Questions of teacher supply and demand, for example, have concerned policymakers through multiple cycles of "shortage" and "surplus;" misgivings about teacher quality are perennial; and the adequacy and form of teacher compensation are matters of continuing public and professional interest, quite apart from the current fascination with career ladders and merit pay. Thus, the point in faulting NCES for the paucity of its teacher data is not that it has responded slowly to the issue of the moment but that it has neglected an area of long-term policy concern.

Current NCES Data on Teachers

NCES today is able to tell us little more about American teachers than how many there are. The CCD surveys obtain annual data on numbers of full-time equivalent elementary and secondary teachers (and other employees) by state and LEA. The state-level data are reported in the Condition and Digest; the LEA-level data go unpublished and, apparently, unanalyzed (see below). NCES collects no information on the composition and characteristics of state or local teaching forces, nor on assignments or working conditions of teachers, nor on salaries or other aspects of compensation. The only data on teacher salaries now published by NCES are estimates of average salary by state reprinted from publications of the National Education Association.[7] The occasional and fragmentary data on teacher characteristics and assignments published by NCES (only for the nation as a whole) are also borrowed from NEA.[8] No information whatsoever is provided by NCES (or NEA) on variations in teacher characteristics or assignments among states or LEAs; nor on aspects of compensation other than the mean teachers' salary in each state.

Aware of these increasingly conspicuous gaps, NCES has recently taken tentative steps to produce teacher data. In a new survey (the data from which are now being processed), teachers in a nationally representative sample of schools were asked to report on their teaching experience, training, assignments, work hours, compensation, and certain personal characteristics.[9] In addition, NCES is about to sponsor an effort to design a new, more extensive survey of characteristics of the teaching force. I understand, however, that neither the just-completed survey nor the projected more extensive survey is designed to produce interjurisdictional comparative data, and so, unless further action is taken, most of the aforementioned data gaps will remain.

---

[7] For example, the Digest, 1983-84 reprints data on average salaries of teachers and instructional staff from National Education Association, Estimates of School Statistics, 1982-83 (pp. 54-56).

[8] For example, the Digest, 1983-84 presents findings from NEA sample surveys on the composition of U.S. teachers by race, sex, highest degree, etc.; on average age and experience; and on average hours and days taught, class sizes, and salaries (Table 43, p. 51).

[9] National Center for Education Statistics, "The Elementary/Secondary Education Data Acquisition Program."

NCES also conducts more specialized surveys aimed at responding to concerns about teacher supply and demand. In a survey of "Teacher Demand and Shortage," LEAs and other educational institutions are asked to report on teaching positions, vacancies, new hires, certification status of teachers, and teacher assignments but not on teacher characteristics or compensation. As I will explain below, this survey is flawed not only because key items are missing but also because it is not based on economically meaningful definitions of "shortage" or "demand." Some information on newly hired teachers and their characteristics, including salary information, is also obtained from a triennial survey of recent college graduates. Neither of these surveys provides interjurisdictional comparative data, however, or brings together the multiple types of data needed for policy analyses. Thus, both are useful only for very limited purposes.

Issues and Information Needs

To show why the aforementioned gaps in NCES teacher data are troubling and what types of additional data are required, I now consider some current and perennial policy concerns and research questions pertaining to teachers and the types of information needed to address them. Specifically, I comment on three topics of current concern: teacher quality, teacher compensation, and teacher demand and supply.

Teacher Quality

Even before the reports of the Commission on Excellence and other reform commissions began to appear in 1983, teacher quality was a major concern of educators and policymakers, and now it has become one of the central foci of efforts to improve the schools. According to the reformers, low teacher quality is at the heart of our educational problems, and drastic changes in teacher compensation, certification, and training are called for. However, little is known about even the most basic quality-related attributes of the teaching force or about the relationships of teacher quality to other factors, and the paucity of NCES data plays a role in preserving this ignorance.

Among the issues potentially illuminable by better data are how the teaching force is and has been changing with respect to certain quality-related attributes; how quality-related characteristics of teachers vary among states, school districts, and schools; how such characteristics relate to teacher compensation, other conditions of teaching, and the state of the teacher market; whether teachers with different characteristics tend to be assigned to different types of schools and pupils; and how teacher attributes relate to pupil achievement and other measures of educational outcome. Without good, disaggregated data on teacher characteristics and assignments, one can do little more than speculate about such concerns.

NCES cannot be faulted for failing to provide data on teacher quality per se, since there is little agreement on how "quality" can or should be measured. On the other hand, there are many teacher attributes, arguably germane to the quality issue, that are not only feasible but relatively easy to measure. A list of some of the more obvious items is as follows:

- Personal characteristics: age, race, sex, languages spoken, etc.,
- Educational background: attainment, fields of study, degrees, institutions or types of institutions attended,
- Experience: years of teaching, assignments, experience in other occupations,
- Career pattern: entry, exit, reentry; rank, promotion, plans and expectations,
- Assignment: grade level, subject, types of pupils, type of school, special programs,
- Work load/working conditions: hours per day (teaching, other), days per year, class size, pupil load, special pupils, support staff, nonteaching responsibilities,
- Compensation: salary, retirement and other fringe benefits.

There are different methods of obtaining such data. The possibilities include sample surveys of individual teachers, surveys of sample districts (or conceivably all districts), and surveys of states. When one adds a longitudinal dimension, the possibilities multiply. As usual, the appropriateness of any given method depends on the purpose for which the data are to be used. I discuss below, under the heading "possible data collection strategies," some of the alternative approaches and their uses. For the moment, I note only the following: first, that the approach toward which NCES now seems to be leaning—collection of data on the characteristics of a nationally representative sample of teachers—is suitable for only a limited range of applications; second, that work on the major issues of teacher quality requires interjurisdictional comparative data on teacher characteristics—something that seems not to be contemplated in current NCES plans.

Teacher Compensation

The system of teacher pay is now receiving an extraordinary amount of attention because of recent reform commission recommendations in favor of merit pay, career ladders, and other forms of teacher incentives. Some states are already acting to install such systems and others are considering the possibility, so there is much interest in assessing existing and alternative salary structures. Even without this special interest, however, teacher compensation is a matter of continual concern to policymakers. Teachers' salaries are the largest single element of education cost, and the level of teachers' pay has long been assumed to be a major determinant of the quality of the teaching force. A valuable side effect of the current debate over teacher incentives has been to make clear how little organized information exists on how teachers are paid in different places and how pay systems are changing and on the consequences thereof for educational costs, the make-up of the teaching force, and ultimately the quality of teaching and educational outcomes.[10]

NCES now collects no information of its own on teacher compensation. Even in the past, before abdicating the responsibility and leaving it to NEA, it collected only average salary data. But average salary figures tell very little about how states or districts pay their teachers. Districts can have identical average salaries but different salary schedules and distributions of experience and training; or, districts can have identical salary schedules but different levels of average pay. It is presumably the salary schedule rather than average pay that determines the attractiveness of a district to teachers and hence the ability of a district to attract quality staff. Thus, even if NCES reinstituted collection of state average salary data, that would hardly suffice to address current concerns about teacher quality and cost.

---

[10] The National Education Association is said to possess a large computer file of individual district salary schedules. Such information could be used to answer a variety of questions about how teachers are paid in different types of school districts and how salaries vary with teacher experience and training. However, the NEA apparently does not use its data base for such analytical purposes or, if it does, does not publish the results.

To see what kinds of data might be worth collecting, consider, first, what it would take to provide reasonably complete answers to the questions, "how are teachers in the United States paid?" and "how does the compensation of teachers vary among school systems and states?" Going beyond undifferentiated salary averages, one would want to know how salaries vary in relation to teacher characteristics, including not only the education and experience factors on which salaries are based but also such other factors as age, sex, race, subject-area specialty, and grade-level assignment. It is more informative, for example, to make interjurisdictional comparisons of the salaries paid teachers with standard qualifications (e.g., a master's degree and five years of experience) than of overall salary averages. Moreover, to go beyond averages in another respect, one would want to know how the teachers in any given category are distributed among salary brackets. For instance, what percentages of high school teachers, or more specifically, of high school mathematics teachers, earn less than $15,000, $15,000 to $20,000, $20,000 to $25,000, and so forth. In addition, to form more complete pictures of teachers' rewards and teacher personnel costs, and to facilitate interjurisdictional and interoccupational comparisons, one would want to broaden the scope of data collection to encompass retirement contributions and other fringe benefits as well as the salary component of compensation. Finally, to describe fully how teachers are paid requires data on salary schedules as well as on levels or distributions of pay—that is, on starting salaries, increments paid for additional units of training and experience, and any incentive features of salary systems.

As to the appropriate unit of analysis, or level of disaggregation, that naturally depends on how the data are to be used. It would be desirable, for example, in conjunction with recent Education Department efforts to compare educational resources and outcomes across states, to produce state-by-state salary data, broken out by some of the categories suggested above. For other purposes, salary data by type of district (i.e., by district size, urbanicity, socioeconomic composition, etc.) would be more appropriate; and for more detailed analyses, data by type of district within each state, or simply data for individual districts, would be needed. In particular, any application requiring information on salary schedules, automatically implies selection of the individual district as the unit of analysis.

Supply and Demand

NCES has attempted, as mentioned earlier, to respond to concerns about teacher supply and demand by conducting its special surveys of "teacher demand and shortage," but there are several respects in which these efforts need to be strengthened. The main shortcomings, in my view, are (1) the lack of economic underpinnings and the consequent omission of categories of information essential to an analysis of demand and supply, and (2) inadequate coverage of the flow of persons into and out of the teaching force.

The absence of a guiding economic conception of supply and demand is most evident from the failure of the survey designers to recognize that supply, demand, surplus, and shortage all have something to do with prices (salaries), and consequently that salary information must be brought together with data on positions, vacancies, new hires, and the like. If there are to be comparative analyses of the supply-demand situations facing different districts or states, differences in salaries offered by those jurisdictions are likely to be key explanatory factors. Thus, the collection of salary information (which, for reasons explained above, means considerably more than information on average salaries) should be integrated with collection of other information bearing on supply-demand issues.

Another important missing element is the concept of a teacher market, or market area. Where many districts seek to hire teachers in the same market (e.g., in a major metropolitan area), the supply of teachers to any single district becomes a function not only of its own salary schedule, working conditions, and other attributes but also of the characteristics of the competing districts. For example, $15,000 per year is a high starting salary in a market where the typical offer is $12,000 but a low one where it is $18,000. To contibute to a better understanding of teacher markets, therefore, NCES should collect data not only on the salaries offered by each jurisdiction but also on how those salaries compare with prevailing salaries in the particular market in question.

A third missing item, related to the two already mentioned, is information on nonteaching salaries in the various teacher markets. Such information is germane because it is relative pay that affects teacher supply. A given teacher pay scale may be highly competitive where nonteaching opportunities for college-educated workers are few but inadequate where such opportunities are plentiful. Analyses of teacher supply and demand, therefore, require data not only on teacher salaries but on salaries in other occupations as well. This does not mean that NCES should get into the business of collecting data on wages and salaries outside education, but it does suggest that such data should be obtained by NCES (e.g., from the Bureau of Labor Statistics) and merged with information on teacher pay.

Turning to the flow of teachers into and out of the system, NCES seems to have recognized the importance of half this phenomenon, the movement of new entrants into teaching, but not the other half, the outflow (turnover) of existing teachers. Moreover, even with respect to the inflow, more information needs to be collected to understand supply-demand relationships and, especially, the changes therein due to changes in teacher standards and compensation. Reporting on the number of new hires, as in the NCES "Demand and Shortage" survey, is not enough. Information should be obtained, in addition, on characteristics of the newly hired teachers, where they come from, and what they are paid. In particular, it is through the hiring process that new state incentive pay plans and tightened certification requirements are likely to have their effects, if any, on teacher quality. Consequently, it would be desirable for NCES to monitor the characteristics of teachers newly hired during the coming years, so that these important policy changes can be assessed.

The teacher turnover/retention process is as important as the entry process in shaping the character of the teaching force, but this is a process about which little is known. To carry out a comprehensive analysis of teacher supply and demand, one would need information on which types of teachers depart with what frequency from districts with different characteristics and different pay and promotion policies. In particular, it is now becoming important to observe how teacher turnover patterns are affected in states that adopt new performance-based pay and promotion plans. These matters are not covered, to my knowledge, by current or projected NCES data collection efforts—an important gap that needs to be filled.

I have learned recently, in the course of an international comparative study of teachers' salaries, that some other countries are able to report routinely on geographically disaggregated flows of personnel into and out of teaching. This would mean, for the United States, reporting each year on (a) the number of persons entering teaching in each state, broken down by source (i.e., new college graduates, teachers transfering from other states, persons shifting from other occupations, etc.), and (b) the number of persons leaving teaching in each state due to death, retirement, involuntary termination, transfers to other states, shifts to other occupations, and departures from the labor force. Such a flow matrix would provide the framework for a wide variety of supply-demand studies. It seems well worthwhile at least to investigate the cost and feasibility of creating such a data set.

Possible Data Collection Strategies

If NCES does become involved in a major way in collection of data on teachers, it will have to make some strategic decisions at the outset. Among these, the most basic concern the choices of units of analysis, respondents, and level of detail. I consider here some of the diverse purposes for which teacher data might be wanted and the degrees to which these purposes might be served by different data collection modes.

One possible objective, clearly of current interest to the Education Department, is to assemble state-by-state data on teachers to add to the comparative displays of state education statistics (the famous "wall charts") distributed by the Department this year and last. The only teacher data now included are pupil-teacher ratios. Other items of potential interest include statewide averages of teacher experience, training, and other characteristics and indicators of the level of teacher compensation in each state, such as the salaries paid, on average, to teachers with specified standard characteristics. Such information could be obtained from state education agencies (which, in some cases, would have to institute new data collection procedures of their own to obtain the information from LEAs); from NCES censuses or, possibly, sample surveys of individual districts; or, in part, from state-representative sample surveys of individual teachers.

Another, much broader objective is to construct a general teacher data base that can be used to support a variety of research and policy inquiries. Such a profile should include information on teacher characteristics, teacher compensation, and the conditions of teaching. Disaggregation to the state level is the minimum required for such a file to be at all useful, and for most research purposes that level of detail would not suffice. For instance, it would be difficult to derive valid conclusions about teacher quality, patterns of compensation, or relationships between teacher characteristics and outcomes without distinguishing, at least, among urban, suburban, and rural districts; districts of different sizes;

and districts of different levels of income or wealth. For in-depth analyses in any of these areas, individual district data would be required. Such data could be obtained through state agencies or from LEAs directly. The choice between the two seems to hinge on (a) the apportionment of the data collection burden and (b) the trade-off between decentralization and data quality. If NCES did choose the direct data collection strategy, it would seem reasonable to take advantage of the main district-level data collection mechanism already in place by appending a detailed set of teacher-related items to the Common Core of Data.

A somewhat more specialized research-oriented objective is to assemble the data needed to address teacher supply and demand issues, including the key issue of how teacher supply, and in particular its quality dimension, responds to changes in compensation and other market conditions. Some aspects of these issues, especially questions on the supply side, can be addressed through sample surveys of individual teachers or college graduates--provided, however, that the samples are drawn not merely to be nationally representative but to allow comparisons among states and types of districts. Other questions, including many on the demand side, require in-depth data from samples of school districts, such as salary schedules and the distributions of teachers upon them. In particular, an analysis of the flow of persons into and out of teaching would seem to require district-level data, specifically including detailed information on those entering and leaving the teaching force.

Finally, a narrower, but currently high-priority objective is to assemble data sets suitable for evaluating the effects of the major changes in teacher pay systems and certification standards now being instituted around the country. This would probably require data from sample districts in states establishing the new systems (i.e., merit pay, career ladders, teacher proficiency examinations, etc.), with special emphasis on data concerning newly hired teachers and teacher turnover. It would also require collection of longitudinal data to determine the effects of the policy changes over time.

This list by no means exhausts the possibilities, but it suffices to make several points. First, geographically disaggregated teacher data are essential for research and policy uses. State-by-state data will serve some purposes, but for many research applications, district-level data will be required. Second, national data, and hence surveys based only on nationally representative samples, are of very limited value. They provide general background information and good numbers to use in speeches but contribute little to understanding how the teacher system works. Third, whatever the unit of analysis and whatever data collection strategy is used, it is important that data on all the relevant aspects of teaching be collected together. That is, data on teacher characteristics, compensation, working conditions, etc. should all be collected from the same respondents at the same times, so that relationships among these variables can be explored. Fourth and last, a nonsubstantive point: different potential uses of teacher data lead to different demands for data, and it is not readily apparent which demands should have precedence. Therefore, if NCES is to enter the field of teacher data collection in a serious way (which I assume it must), it should first engage in a series of priority-setting exercises and feasibility studies to produce a coherent plan. I would hope that such a review could begin with as blank a slate as possible—i.e., without preconceptions regarding the modes of data collection or the continuation of current or projected NCES surveys in their present forms.

STATE-LOCAL STATISTICAL REPORTS AS A NATIONAL (AND NCES) RESOURCE

In closing, I offer one additional observation that cuts across the areas of school finance and teachers (and probably other areas of NCES data collection as well). I have long believed that a great deal of useful data for educational research and policy studies is produced by states (and, in some cases, by local districts) but that the lack of an organized, central collection greatly limits its use. I suggest, therefore, that NCES take on the role of bringing such material together in a well-maintained central repository.

In the finance area, virtually every state produces an annual report on the finances of its school districts. These reports typically present district-by-district expenditure and revenue data in substantial detail and also cover tax bases and tax rates. The level of detail is often considerably greater than NCES could be expected to collect. Even though such data are often not comparable among all states, there are some cases in which such data could be used to supplement and fill gaps in NCES' own data bases, and there are many more cases in which data for selected states would serve the purposes of research and policy studies. By systematically acquiring the annual reports of all the states, therefore, NCES would be able to create a valuable, multipurpose analytical tool. In addition, much useful information could be extracted from the budgetary documents and financial reports of local districts, selected samples of which (e.g., from large districts) should be included in such a collection.

In the teacher area, state reports are more variable in coverage and format, but many states do produce reports on numbers of teachers (and other staff), certain staff characteristics, and staff salaries by district. Some of these reports are extremely detailed. California, for instance, has produced for many years an annual volume providing not only the full salary schedule of each district in the state but also the number of teachers in every cell of that schedule. In some cases, official state documents on teachers and their salaries are supplemented by reports produced by teachers' unions, containing data on salaries, salary schedules, and working conditions in the various districts. Like the state-local data on finances, the data on teachers, if brought togethr, could satisfy a wide variety of data needs. In addition, during the several years it will undoubtedly take NCES to produce new teacher data of its own, compilations of selected teacher data from state sources might fill some of the currently unmet information needs.

Establishing a central data repository seems a natural role for the national education statistics agency. The availability of such a resource in-house might also have the beneficial side effects of keeping NCES staff in closer touch with developments in the states and providing means of cross-checking NCES' own data. At least a preliminary inquiry seems warranted to determine what data are regularly produced by the states, what uses they might serve, and how much it would cost NCES to develop and maintain such a collection.

EDUCATION AND EMPLOYMENT: SUBSTITUTION POSSIBILITIES
AND
THE TEACHER LABOR FORCE: SUPPLY AND DEMAND

Dr. Sue E. Berryman
The Rand Corporation
Washington, D.C.

June 21, 1985

This paper was prepared for the Elementary and Secondary Data Redesign Project of the National Center for Education Statistics.

SUBSTITUTION POSSIBILITIES IN EDUCATION AND EMPLOYMENT, OR: IT MATTERS IF THERE IS MORE THAN ONE ROAD TO MANDALAY

INTRODUCTION

Key to answering an astonishing number of policy questions is understanding two relationships. *What is the relationship between alternative learning opportunities and the acquisition of skills?* Specifically, what learning opportunities--different types of formal education, leisure activities, and work experiences--are interchangeable with each other in producing skills, and under what conditions?

And, *what is the relationship between the nature of skills and access to occupations?* How does the array of skills that individuals possess map onto the array of occupations, as evidenced by employment in the occupation? How do tight versus loose labor markets for the occupation change the mapping? How does the nature of the employer's job structure affect the mapping--for example, a structure with fairly rigid, narrowly-defined job categories versus one with broadly defined categories? How does the "age" of the occupation (emerging or established) affect the mapping?

The National Center for Education Statistics (NCES) and other agencies now collect data that allow limited analytic progress with these questions. However, as discussed in more detail later, current longitudinal data bases do not approach their potential for supporting research into these issues.

Answers to these two questions can illumine a surprising range of policy issues. I discuss three examples here: shortages and oversupplies of skills; the contribution of formal corporate training to the human capital of the labor force; and theories of labor markets that underlie federal training policies for disadvantaged individuals and dislocated workers.

ROLE OF INTERCHANGEABILITY IN EDUCATION AND EMPLOYMENT POLICY

Skill Shortages and Oversupplies

Policy concerns about oversupplies or shortages of particular skills--scientists, engineers, or public school teachers, for example--are hardy perennials. These concerns, often escalated to statements about "crises," include concerns not just about shortages or oversupplies of skills, but also of those who produce them--for example, of mathematics and science teachers at the secondary level and science and engineering faculties at the post-secondary level.

These "crises" presume that shortages or oversupplies matter. However, we lack data and analyses to determine if and when they matter, for whom, and in what way. For example, do shortages of computer science teachers at the post-secondary level affect the supply of computer skills that employers need, and, if so, for all or for specialized skills only? Are there alternative sources of instruction? In the computer case we know that students in all of the highly quantitative fields gain experience with computer-based equipment and software. In fact, one of the explanations of the heterogeneous backgrounds of those employed in computer science jobs[1] is that opportunities to learn the skills required for these jobs are widely diffused among academic departments.

How do employers and individuals trained in an oversupplied field respond to oversupplies, and what costs do these responses exact? In other words, *do we need to worry about oversupplies?* Which subgroups tend to absorb the costs of oversupplies? Those newly trained in the occupation? Those approaching retirement? How does the accommodation occur? Under what conditions do employers offer and individuals accept lower wages? When do employers ratchet up the education that they require of new hires without changing the skill content of the job? Are individuals who accept such jobs underemployed, or do they tend to "upgrade" the work performed to match their skills? How often do trained individuals leave their field of training? Do they tend to enter fields that use some part of their past educational and work experience? Are these individuals permanently lost to the field in which they were trained, or do they tend to migrate back into it as supply and demand for the field equilibrate?

How do employers respond to shortages of individuals trained in a specific field? In terms of educational and work backgrounds, how *wide* a net do employers seem willing to cast to fill positions? How low in the *level* of training in the shortage field (e.g., ratcheting down from an M.S. degree to an B.S degree) do employers seem willing to go? What are the consequences of staffing jobs with less advanced or less germane training? How "tolerant" are different occupations of variations in skills, i.e., how wide a spread of skills can they accommodate before productivity declines? How do discrepancies between the skills required and those hired become resolved? Do occupants of jobs learn the skills required or do they transform the job to fit their skills?

A study of how employers staffed three electronic data processing (EDP) occupations (computer operators, programmers, and systems analysts) between 1965 and 1970 illustrates employer responses to shortages in an *emerging* occupation.[2] The authors found that a third of the 1970 EDP labor force had been employed in EDP occupations in 1965; a third had entered the labor force since 1965; and a third had worked in non-EDP occupations in 1965. The majority of the lateral transfers did not come from disciplines and occupations related to EDP occupations; only 30.2 percent of these transfers had been engineers, mathematicians, life and physical scientists, engineering and science technicians, accountants, bookkeepers, and office machine operators. Thus, in the EDP case, workers in *unrelated* fields constituted an important part of the labor response to rapidly growing demand.

---

[1] Table B-8, National Science Foundation, *Characteristics of Recent Science/Engineering Graduates:1982,* Report No. NSF84-318, pp.23-24, U.S. Government Printing Office, Washington, D.C., 1984.

[2] Sheldon Haber and Robert Goldfarb. "Labor Market Responses for Computer Occupations," *Industrial Relations,* 17 (1), 53-63, February, 1978.

The data on supply sources showed that entry into the more complex EDP fields, such as systems analyst, required more occupationally-germane formal or on-the-job training. In general, however, the data showed that formal education (e.g., a college degree) could substitute for work experience, that formal education could substitute for vocational training, and that less experienced workers needed more formal education if they lack vocational training.

The EDP occupations were not established occupations in the 1960's, and employers' responses to shortages in these occupations may be a special case, i.e., not germane to how employers handle shortages. Higher skill occupations that are new are not well integrated into educational and employment systems. They lack generally understood entry paths (training sequences, school-work transitions, career ladders) and developed training programs. Their job content tends to be unstable, making it difficult to design relevant educational programs. Thus, de facto, they are higher skill occupations with low entry barriers-- "bright," i.e., "trainable," individuals can enter them relatively freely.

Corporate Training and Human Capital

Although we know more about formal corporate training today than we did ten years ago, we still lack systematic information about which employers provide what kinds and how much training to what kinds of employees for what purposes and with what effect. Thus, the relationship of corporate training to other sources of education and training and its contribution to the human capital of the labor force are not well understood. Does corporate training complement or duplicate other training sources? Total corporate investments in training may be large. However, does a large dollar total translate into substantial changes in human capital? Or does corporate training augment human capital only marginally--either in the sense that the training time *per employee* is so short that little could be learned or the substance of the training has minimal transferability (e.g., company safety procedures)?

Not knowing what role corporations play in creating human capital is like misspecifying a model. Or, using another analogy, it means working with a map of the education and training system that contains a sizeable *terra incognita*. Maps of this kind can produce at least two kinds of diagnostic errors. Those using the map can assume that the unknown domain is "picking up" training not performed in the known parts when in fact it is not. Or they may simply ignore the unknown part. In this case they will underestimate the amount and types of training being conducted, potentially funding training that simply duplicates what employers are already buying.

Federal Training Programs and Images of Labor Markets

The federal government has invested billions of dollars in training and employment programs for various disadvantaged groups. These programs are predicated on (often unspecified) assumptions about relationships between skills and jobs. Are these images of the labor market correct?

In the early 1980's computer-based changes in the economy's technological base and realignments of international markets combined with a recession to displace a large number of experienced workers. As in the early 1960's, these events generated considerable policy discussion about retraining programs. This solution presumes that the problem is a lack of skills for available jobs, rather than a lack of jobs (regardless of the available skills), or low wages for available jobs for which dislocated workers qualify. Since fragmentary evidence indicates that individuals move fairly freely among less skilled jobs *apparently* without additional training, the dislocation problem is not necessarily a retraining problem.

Conclusion

Key to answering these and other questions is understanding interchangeability among sources of skills and how skills map onto occupations. In the United States we tend to assume that a particular occupation requires a unique bundle of skills, obtained from limited sources. This assumption affects curricular design and feeds, if it does not produce, anxieties about skill shortages and oversupplies. If individuals can only obtain required training from highly limited sources, they cannot adjust to shortages in the supply of these sources. If occupations require distinct bundles of skills, they operate like countries with trade barriers. Goods cannot move freely among nations--or, analogously, labor cannot move freely among occupations to adjust to changes in demand.

In fact, fragmentary but internally consistent empirical evidence indicates that this assumption reflects a mechanistic and fundamentally inaccurate view of human cognitive capacities, the educational system, the effects of schooling, and the nature of jobs. Substitution possibilities pervade the educational system and the work place.

At least some of the same skills can be obtained in alternative ways--in different courses, from different schools, during military tours, in civilian jobs, or from volunteer or avocational activities. Substitute sources of knowledge can vary in degree and number, depending on the knowledge in question. Two alternatives may be complete substitutes, i.e., virtually indistinguishable from each other, or only partial.

In the labor force different jobs can be filled with individuals who have more or less heterogeneous educational and occupational backgrounds. Again the substitutes vary in degree and number, depending on the job in question. Fragmentary data suggest considerably more interchangeability, even for high skill jobs, than we commonly assume, especially for the newer occupations that technological innovations tend to spawn. The substitution possibilities, not surprisingly, seem greater for less demanding skills, such as computer programmer, than for more demanding ones, such as computer systems analyst. However, the data show that students move between post-secondary fields of study and that even highly skilled workers move between high skill occupations. The movement is not unconstrained, but there is movement.

DATA REQUIREMENTS

NCES is not the appropriate organization to collect the data required to answer some of the questions posed here. For example, laboratory experiments or observational studies are probably preferred to survey techniques for studying the processes by which employees accommodate discrepancies between job skills required and those possessed.

However, the NCES and other federal agencies such as the Department of Labor and the National Science Foundation support certain data collection activities that could support progress on substitution questions. These are longitudinal data bases that measure characteristics of workers' formal education and training, their work experiences, and their movement across time among occupations, firms, industries, and geographic areas.

Existing longitudinal data bases--two of them funded by the NCES--contain some information on the critical variables: for example, *The National Longitudinal Study of the High School Class of 1972*; *High School and Beyond*; *The National Longitudinal Surveys of Young Men, Young Women, Mature Women, and Mature Men*; *The National Longitudinal Survey of Youth Labor Market Behavior*; *Panel on Income Dynamics*; and *Experienced Scientists and Engineers*. However, these data bases all lack the detailed *taxonomies* of educational and work experiences required for completely satisfactory analyses of substitution issues. Carefully developed taxonomies are critical to the success of these measures. Examples of taxonomies that would have to be developed are:

- the *sources* of formal training, such as vocational training in an area vocational high school, vocational training in a comprehensive high school, the military, an avocational activity with some kind of formal instruction, corporate training delivered outside of formal educational institutions, corporate training that uses formal educational institutions (by type of institution), a public two-year college, a post-secondary proprietary vocational school, a four-year college, a university, a corporate college.

- the *subjects* of formal instruction, such as particular courses in mathematics or computer science, a course in using a company's text processor software, a finance course.

- the amount of formal training, such as two versus four years of high school mathematics, a two year associate degree program versus a four year B.S. degree, a four year versus a five year B.S. program in engineering, a week versus six weeks of management training.

- *characteristics* of the work experience that measure development and deployment of skills: for example, the Census 3 digit occupation, the technology (-gies) used in the job, the nature of supervisory responsibilities, time diaries--similar to those used to measure work in the home--to measure routine and periodic work activities.

KNOWLEDGE ABOUT THE NATION'S TEACHERS, OR: YOU'VE LOST THE WAR IF YOU CAN'T FIND THE BATTLEFIELD[1]

INTRODUCTION

In 1982-83 the nation's public elementary and secondary schools employed over 2 million classroom teachers. This labor force is almost exactly the same size as the nation's active duty enlisted force and officer corps. In terms of the number employed, the elementary/secondary teaching occupation is the largest professional and technical occupation and among the five largest white collar, blue collar, and service occupations.

The public and private cost of creating and employing this labor force is enormous. Last year the nation spent over *one hundred billion* dollars on public elementary and secondary schools, over fifty billion of these dollars going to teacher salaries. Since almost all public school teachers have at least a bachelor's degree and half have at least a master's degree or 6 years of post-secondary education, the size of this labor force represents an enormous public and private investment in human capital.

However, despite the tremendous cost of creating and employing public school teachers, we know almost nothing about how many teachers we will need, when, at what levels, in which disciplinary fields, and in which parts of the nation. We do not know if the publicized shortages of mathematics and science teachers are general shortages or spot shortages. Current efforts to improve supply and demand projections are being severely limited because we lack the fundamental information required to predict supply and demand.

More important than the costs of creating or paying it, this labor force develops a substantial part of the human capital in each new generation of children. Those 25-29 years old in 1980 had completed a quarter of a million years of education, *89 percent of these years at the elementary and secondary levels*. Elementary and secondary public school teachers particularly affect poor children: these children are more apt than non-poor children to receive their education in public schools, they are less apt to have well-educated parents and thus rely more heavily on teachers for their academic training, and a larger proportion of their total years of education occur in elementary and secondary schools.

However, despite the consequences of teachers' job performance for the human capital of the country--especially for that of certain subgroups, we have no acceptable data on the quality of new or experienced teachers. For example, we know that the decline in SAT scores of new teachers in the last decade is greater than that for the total population of SAT test-takers, but we do not know the relationship between score declines and teaching performance. We have no way of assessing the truth of the speculation that schools are adjusting to presumed shortages of teachers by resorting to poorly qualified individuals.

---

[1] This section is based on work conducted for Edward J. Meade, Jr., of The Ford Foundation.

REFORM OF THE TEACHING FORCE

A blizzard of recent reports on American schools calls for reform in our public schools to improve the quality of the education that our children receive. These range from *A Nation at Risk,* issued by the National Commission on Excellence in Education, to the study conducted by the Carnegie Foundation for the Advancement of Teaching, *High School.* These reports identify our teachers as part of the problem and potentially as part of the solution.

Partly in response to these reports, state after state has legislated or is contemplating legislation that will affect the teaching force. These bills will not only affect the requirements for teachers, but also the flows of individuals into and out of teaching. Thus, ultimately, they will affect the stock of teachers--their number, field of teaching expertise, and quality. These reforms range from salary increases to changes in high school graduation requirements to competency tests for teachers.

This flurry of legislative activity reminds one of the blind man and the elephant. Educational reformers have hold of different parts of the animal--the tail, an ear, or a tusk, and each believes it to be the whole animal. In fact, the teaching labor force is a dynamic, human resource system. At any point in time, the stock of teachers--their numbers, their specialized skills, and their quality--reflect numerous prior decisions by individuals. These choices include the decision to train as a teacher or not to train; to train as a high school teacher or as an elementary school teacher; to train as a bilingual or as a foreign language teacher; to enter teaching or not to enter teaching; to stay in teaching or to leave it for home responsibilities or another occupation in the labor force; to stay in teaching or to retire; to stay in the same school or to move to another school, another district, or another state; to re-enter teaching or not to re-enter teaching. These choices are not random, any more than choosing to enter, stay in, or leave any occupation--lawyer, secretary, manager, chemist--is random. They reflect the relative attractiveness of the occupational alternatives available to the individual. As teaching becomes less attractive relative to these alternatives, the number and quality of individuals who elect to enter and to stay in teaching declines.

Educational reforms, undertaken for whatever reason, will intersect with individual choices and change the teaching force in many ways. However, *when this labor force is not conceived of as a system, when the data do not exist to diagnose problems with this system at appropriate policy-making levels (national, state, or district), and when the data do not exist to monitor, let alone project, how particular reforms* *affect the behaviors of potential, new, and experienced teachers, we are flying blind.* Under these conditions the chances of misdiagnosing problems are high; the chances that reforms will produce the changes that reformers want are low. As we learned so bitterly in the 1970's, failed social reforms carry very heavy costs. The greatest is that the problems are not alleviated, while the fickle political mood that created the opportunity for reform has dissipated. Cynicism about the responsiveness of schools and teachers to national concerns deepens, and the political will to "try again" lessens. Finally, reforms, whether successful or unsuccessful, perturb the system, and turbulence is costly.

DATA ON THE TEACHING FORCE

Thus, while we need action, it needs to be informed action. However, the state of the data required for informed action is, quite simply, shocking. This is especially true when we compare our information about teachers with that which we have on labor forces of comparable size, national importance, or human capital, such as the enlisted armed forces or the scientific and engineering labor force.

Data Needs

The questions that the public and educational policymakers ask about the teaching force indicate that they want information on at least six dimensions of supply and demand. These are:

1. the quantity, or *number,* of teachers;

2. the *quality* of teachers;

3. the number and quality by *level,* a level being defined as a grade category (such as the grades 4-6) that requires at some teacher skills not required by other categories;

4. the number, quality, and level by *field,* field being defined as the teacher's special field of training, such as mathematics, bilingual education, biology, English, or art;

5. the number, quality, level, and field of teachers by their race and ethnicity; and

6. the number, quality, level, field, and race and ethnicity of teachers by their *geographic distribution,* the geographic units of interest being districts, states, and the nation and unique supply and demand environments, such as inner city schools or rural schools.

The reasons for assessing the number and quality of teachers are self-evident. The level and field distinctions presume that teachers are not entirely interchangeable with each other. The extent to which teachers with different educational and experiential backgrounds can properly substitute for each other in the classroom is an empirical, critical, and unresolved question. However, we can safely assume that not all levels or fields are interchangeable with each other.

The racial and ethnic characteristics of the teaching force are politically, perhaps pedagogically, and, in some states, legally important. Although it is easier to measure the racial and ethnic characteristics of supply than of demand, court orders define demand for districts legally required to balance the racial and ethnic composition of students and faculty. These tend to be Southern districts. In other districts the racial and ethnic composition of the students probably establishes a demand "range" for teachers of a particular race or ethnicity.

The geographic dimension is extremely important. The nation's public schools vary greatly, and the possibilities for quality variations, geographically-specific teacher supply and demand imbalances, and resultant educational inequalities among school districts and states are enormous.

Data Sources

There are four fundamentally different sources of data on teachers. One source is organizations directly responsible for the delivery of public elementary and secondary education. These include schools and local boards of education. The second source is public administrative and legislative groups at the state level, such as state education agencies, state boards of education, and education committees in state legislatures.

The third source is associations--groups whose membership consists of particular categories of education practitioners and policymakers. These include teachers' organizations, such as the National Education Association and the American Federation of Teachers; associations of teachers of particular subjects, such as the National Science Teachers Association, National Council of Teachers of English, and the American Vocational Association; chief state school officers (the Council of Chief State School Officers); and associations that represent schools in particular environments, such as the Council for Great City Schools.

The fourth source is organizations whose primary or sole function is data collection or analysis and whose substantive focus includes education. These groups, fundamentally non-partisan, may be public or non-profit. They include organizations such as the National Center for Education Statistics, the Bureau of Labor Statistics, the Bureau of the presumed to imply quantity considerations only. However, the Census, the Rand Corporation, the Urban Institute, Educational Testing Service, and the National Center for Education Information.

Status of Data on the Teaching Force

State and local data. Interviews with the first two sources (state and local organizations) and with groups in the third and fourth categories about data from the first two sources indicate the following problems with data collected by schools districts and the states.

- In general, teacher data at the district and state levels are, in the words of one knowledgeable respondent, "deplorable," "spotty," and "inconsistent."

- All states collect data from their districts on the number of student enrollments and the number of teachers. These data are usually--but not always--collected by level (elementary versus secondary). They also usually have data on education revenues and expenditures by categories, but use such different definitions of what constitutes revenues and expenditures that these data are usually non-comparable across states.

- A few states, such as Florida, Illinois, and New York, routinely collect other data about teachers. For example, Illinois and Florida conduct supply and demand studies for their states and collect the data required for fairly simple models. However, data such as these are usually not comparable across states because the definitions of variables differ.

- Our interviews suggest that districts often have data about teachers. Since these data are almost never used for analytic purposes, we do not know how district data vary in their accessibility, quality, comparability, or temporal and substantive coverage. The few cases where researchers have tried to use district data about teachers suggest that, although the data exist, it is a major task to get them to the point where they can be used to answer questions about the district's teaching force. The data are scattered or not compiled. Our experience with data of any kind collected at local levels strongly suggests that district-level data will not be comparable across districts. It also suggests that the quality and temporal and substantive coverage will be spotty.

We tried to identify the conditions that encourage states and districts to collect teacher data. Our respondent interviews stunningly revealed why good data about teachers are not routinely collected by states and districts. *In most states the key actors do not conceive of public education as a major enterprise whose efficiency and quality might be improved if they knew more about it*--for example, if they knew more about inputs (such as teachers or laboratory facilities), tradeoffs between inputs (such as more teachers versus modernized laboratories), and about how inputs get used. As a result, they do not see the need for teacher data.

Essentially, districts and states only collect data in response to specific incentives--a behavior pattern that produces the "spotty" data collection on which so many have commented. For example, most districts collect data on enrollments or the number of handicapped students because state or federal aid is contingent on them. Other data are collected in response to mandates from state legislatures, state boards of education, or state departments of education. These mandates spring from particular interests of legislators or administrators or from national concerns about the health of education that find voice at local and state levels.

These state and district behaviors are not surprising. States do not collect data for several reasons, the philosophically and politically most important being that Americans see education as primarily a local prerogative and responsibility. Although the balance of power between the state and local levels varies within a state and from state to state, public education is still generally locally controlled. Since policymakers at the state level usually have very limited power to act on any data that they might collect, they tend not to collect much.

Other reasons that states do not collect teacher data are often related to the same issue--the distribution of power among the local, state, and federal levels of government. These include state staffs too small or not competent to collect data, no pressure to collect data ("no mandate"), or the fear that data will make their educational system look weak. States stop collecting data that they once collected because the actors change--those responsible for mandating data collection leave office, or because policy questions that led to the initial collection of data have been answered.

At the district level, there are *economies of scale problems* that reduce *district* incentives to collect and analyse teacher data. Quality data bases on teachers require substantial resources, especially at the front end. Essentially, although some districts employ thousands of teachers, most districts are small employers. Empirical studies show that firms have very different investment behaviors, depending on their size. For example, small companies are much less apt to run formal training programs for their employees.

These studies suggest that we are more likely to find good teacher data in the very large school systems than in the smaller ones. We suspect that, implicitly, administrators of smaller districts tend to think that they would not get enough return to warrant the cost of designing and maintaining a data base. For example, they may make some kinds of decisions only infrequently. Or the number of individuals affected by any given decision is small. Or the district is so small that administrators know what is happening through personal contact and can keep the relevant data in their heads. Or as one of often hundreds of districts in the state--and of thousands in the nation, the district may have limited control over issues that they care about, thus reducing the incentives to collect data germane to these issues.

Data from associations and statistical and analytic groups. We reviewed what data exist from these sources for estimating the six dimensions of current and projected teacher supply and demand. This review revealed the following data problems.

1. *We have no data from any source on certain variables.* There are no data that can be used in any serious way to assess the quality of new or experienced teachers. We have no data on the geographic mobility patterns of *new* teachers--where they look for jobs and how far and where they move to take a job. Although we have old data on the rates at which experienced teachers change schools, districts, or states in which they are teaching, we lack even old data for converting these aggregate rates into flows between specific states and districts.

We have no data on teacher benefit packages. Thus, we do not know what is in these packages, how they vary, or how transferable components such as retirement benefits are across school districts and states. Any serious estimate of the total compensation package for teachers needs to include non-wage benefits. Although we have data on whether districts offer special monetary incentives, we do not know amounts or probabilities of receiving these payments. The amounts or chances of receiving them may be so small that they can be discounted as incentives and ignored in estimating compensation packages.

2. *We have only very old data.* We have only old data for some variables, but changes in factors that affect teacher supply and demand, such as changes in the relative attractiveness of teaching, make these data obsolete. For example, data on teacher turnover are 15 years old.

3. *We have no time series on the variable.* One of the best ways to assess the effect of a new policy is to look for changes in the variable in question *after* the introduction of the policy. These assessments require measures of the variable before, as well as after, the new policy is introduced--and preferably measures at several time points before the policy change. We rarely have comparable measures of a variable across time.

4. *We have data, but from different sampling frames, differently worded questions, and different time periods.* These non-comparabilities make it impossible to pool available data to increase the statistical precision of "snapshot" (cross-sectional) estimates or to create a time series.

5. *We have only partial data.* This problem arises for most variables of interest. We will have data on some issue, such as the age structure of the teaching force or salaries, but lack these data by important distinctions such as level (elementary or secondary, for example), field (such as mathematics or English), or for different administrative units (nation, state, district, school). There is tremendous, policy-relevant heterogeneity in the teaching system, and having data for the total force without these distinctions renders them almost useless.

6. *The sources of data are not credible.* When groups with political interests collect the data, even if these data are entirely valid, they remain suspect. The National Education Association is the only source for many data on teachers--salaries, for example, and for several variables measured in their survey of the *Status of the American Public School Teacher,* such as working conditions and job satisfaction. These data seem professionally collected and reported. However, the NEA is a stakeholder in policy debates, and their data are not necessarily seen as trustworthy.

7. *The measures have validity problems.* Important measures whose validity is questioned are the measures of teacher shortages and out-of-field teaching. Recent national survey data on these variables do not indicate serious shortages even in the fields of mathematics and science; they show low rates of out-of-field teaching. These data do not fit what many think is happening in schools. A source of the discrepancy between data and impression lies in the definitions of the key concepts. For example, these data refer to "shortage" defined as unfilled job vacancies. However, what many mean by "shortage" is an inadequate supply of *appropriately trained and experienced* teachers. Data will not be trusted unless those who need them agree on the measurement system behind the data--on the definition of the concept to be measured and on its operationalization, as embodied in a specific measure.

8. *Required data may exist, but their potential for supply and demand analyses of the teacher labor force has not been assessed.* For example, the Bureau of Labor Statistics of the Department of Labor collects data on separation rates by occupation and reason (e.g., retirement). Can these data be used to estimate teacher turnover rates and post-exit destinations? What are their limits? The Census Bureau collects data on geographic mobility by occupation. Can these data be used to estimate flows of teachers in and out of geographic areas?

9. *The current data system seems inefficient.* For example, the National Center for Education Statistics, Equal Employment Opportunity Commission, and Office for Civil Rights send separate surveys to the same respondents. A single, longer survey would eliminate non-comparabilities introduced by different agency procedures and different data collection times. A consolidated survey should also reduce the burden on respondents. In other cases the same data seem collected at state and district levels, but data collected from districts can be aggregated to yield state estimates.

THE ROLE FOR THE NATIONAL CENTER FOR EDUCATION STATISTICS

In this section we describe why we think the the primary responsibility, authority, and resources for collecting data on the nation's teaching force should reside in the NCES. We also describe roles other organizations might play in conjunction with the NCES.

Why Vest Data Responsibility in the NCES?

The National Center for Education Statistics, or the NCES, is the nation's federal data center on education questions--public and private; elementary, secondary, and post-secondary; student enrollments, dropouts, and graduates; revenues and expenditures of public schools;. institutions and teachers.

. The function performed by the NCES has been performed by the federal government since the nineteenth century, although this function has been vested in differently organized units with different names. In fact, the main rationale for the old Office of Education and its organizational predecessors was the collection and dissemination of statistics on education. Although NCES currently lacks a cogent vision of what data should be collected about teachers, it operates in the non-political and professional tradition of the U.S. Census Bureau and the U.S. Bureau of Labor Statistics (BLS). Just as the politically sensitive monthly unemployment figures released by the BLS are considered independent of the political party in power, so NCES data are considered independent of the many stakeholders in education.

Thus, the NCES meets three criteria for an adequate data system about teachers. First, positioned at the federal level, it is formally authorized to work with all of the nation's school districts and states. Second, it receives an annual Congressional appropriation of funds to be used for the express purpose of collecting data about education, including data about teachers.

Finally and perhaps most important, it is independent of the many stakeholders in education. To be used, data have to be trusted. Their collection and analysis must be and *must be seen to be* non-partisan. If a group with a political stake collects the data, no matter how valid the data are, they will be perceived as compromised.

By virtue of its federal position, the NCES is independent of any state or district. It is independent of the many associations in education. It is subject to the balance of power that affects all federal agencies. Finally, it is included in the special agreement struck between the Congress and the Executive Branch about the importance of non-partisan data on issues of national import. This agreement, although it has to be defended periodically, has created an independent, non-partisan status for the federal statistical agencies.

It can be argued that since states and school districts make most of the decisions about the educational system, they, and not a federal agency, are the logical units for collecting data on teachers. We argue that there are several compelling reasons to keep the basic data collection at the federal level. First, for reasons specified earlier, most states and districts have historically collected only minimum information about the elementary and secondary public school system, including teachers. We have no reason to think that the conditions that discourage data collection at these levels will permanently change in all states.

Second, many public education issues require data comparable across states. The issue may involve geographic units larger than an individual state--for example, estimates for the nation or for a category of states, such as the sunbelt states. Or it may require the ability to differentiate general from state-specific problems.

For example, many issues, while not federal, are framed as national issues, i.e., true for all states. Most recently, these include the quality of secondary education and shortages of mathematics and science teachers. Without comparable district and state-level data, national study commissions or federal policymakers cannot distinguish general from dramatic, but localized, problems. The legislators of a specific state or the board of education of a specific community cannot determine when something, while generally true for the country, is not true for them.

Thus, if states or districts are the sources of teacher data, *the data have to be comparable across states or districts*. Although in theory procedures could be established that would yield comparable state or district data, in practice this becomes a formidable task for more than a *very* few standard variables.

Third, states and districts cannot obtain data on certain key teacher supply and demand questions. These are questions that involve *relationships* between states--for example, questions about the geographic mobility of newly trained and experienced teachers--and individuals not employed by a state's school district--for example, questions about factors that affect undergraduates' choices of fields of study or the decisions of those in the reserve teaching pool to enter (re-enter) teaching.

Fourth, states and districts are not necessarily the best agents for collecting sensitive data from teachers in district or state employ. For example, teachers may be reluctant to report their job satisfaction, retirement plans, or plans to change occupations if the data are collected by their employers.

Role for Other Groups

An adequate data system about teachers must provide the data that all of the nation's major users require. These users include those with public responsibilities for elementary and secondary education and large interest groups with legitimate political objectives. In other words, they include users at different levels of government--national, state, and local, such as state or county boards of education, state and local education agencies, national and state legislators, and the U.S. Department of Education. They also include interest groups, such as representatives of minority groups, partisan state legislators, and representatives of the teachers' organizations.

Thus, although we argue that the primary responsibility, authority, and resources for collecting data on teachers be lodged in the NCES, *designing* that system requires the participation of the major users of those data. A participatory process is a slow process, but it will increase the utility and political fairness of the ultimate data system. It will also increase the chances that state, district, and association groups will co-operate with the NCES in implementing the system.

Supplementary Documents

Committee on National Statistics, National Academy of Sciences. "Preliminary Compilation of Descriptions of National Data Sets Relating to Supply of and Demand for Precollege Science and Mathematics Teachers." Unpublished paper. Washington, D.C., August, 1984.

Reisman, A. et al. "On the Voids in U.S. National Education Statistics." Unpublished paper. Cleveland, Ohio: Case Western Reserve, February, 1985.

Stern, Joyce D., et al. "Report of the Planning Group on Teaching and Teachers." Unpublished paper. Washington, D.C.: The National Center for Education Statistics, February, 1985.

DATA COLLECTION FOR IMPROVING ELEMENTARY/SECONDARY EDUCATION

John H. Bishop

Associate Director--Research Division

The National Center for Research in Vocational Education
The Ohio State University

I. INTRODUCTION AND SUMMARY

Since education is a local responsibility, the purpose of federal data collection is to help the citizenry and political leadership of individual states and localities to make better decisions. Better decision making at the local level requires two things (1) better understanding of how schools influence learning and how effective schools develop and evolve and (2) comparative data on the performance of state and local educational systems so the citizenry and public officials can hold local school administrators accountable. The collection and publication of data is a public service that benefits everyone in the country and thus is properly a federal function.

Achieving the first goal--better understanding of how to improve education--requires large longitudinal data sets like the National Education Longitudinal Study of 1988 (NELS88) which follow students through their schooling and into the labor market. What is needed is large data sets which merge many different kinds of data:

- achievement tests for a great variety of subjects (not just one subject at a time) which have separate subtests for higher order skills and basic skills
- aptitude tests
- school records on courses taken, grades, absenteeism, discriptions of special education service received, and test scores
- questionnaires measuring student background and attitudes (i.e., locus of control)
- parent questionnaires
- surveys of the students, teachers, guidance counselors and principals
- multiple administrations of tests and surveys about 2 to 4 years apart
- follow-up data on labor market outcomes with complete histories of employment and unemployment up to age 40
- interviews with employers to measure skills and job performance of recent school leaners.

The public needs to be able to monitor the progress and achievements of their state and local systems of education. The logical point of comparison is the progress and achievements of other states so the primary focus of NCES's core data program should be on providing data that is comparable across states. Emphasis should be placed on measuring outcomes rather than inputs. The key data elements that need to be added to the program are the following:

- Performance on achievement tests in a variety of subjects that all (or almost all) students in the school take. (It is feasible to put different tests on a common metric and the effort should be made.)
- Comparisons of a states relative performance in tests for first graders with performance in later grades. (This is a better measure of educational effectiveness than mean test scores which are heavily effected by the social background of the student body.
- Separate test scores for basic and high level skills.
- State comparisons of the economic outcomes of schooling: earning, employment, occupation.
- Time series and regional data on salary offers to recipients of associates degrees by area of study.
- Time series and regional data on the economic outcomes of schooling from an improved Current Population Survey.
- Time series and regional data on how schools influence the development of character (e.g., locus of control).
- International comparative data on achievement in particular subjects and on time use.

II. NEED FOR DATA THAT ACCURATELY COMPARES STATES AND LOCAL EDUCATION AGENCIES

Responsibility for and administration of elementary/secondary education lies in state, local and private hands. It is important to remember this fact when planning a federally sponsored system of data collection on elementary and secondary education. The primary purpose of the common core data program is, therefore, not to help the federal government make better decisions but to help the citizenry and political leadership of individual states monitor the progress and achievements of the educational system in that state. Progress and achievements can be defined only when some standard of comparison is available. Since the natural point of comparison is the progress and achievements of other states, the key requirement for the common core data program should be the accuracy of the across state comparisons of the levels (e.g., expenditure per pupil in average daily attendance) or changes (e.g., changes in performances on standardized tests) in various indicators of effort and performance.

In order to effectively manage their responsibilities on a day-to-day basis, state governments and local educational agencies (LEAs) need timely and accurate data on enrollment, staff, revenues, and expenditures for each of the schools and school districts under their charge. From the point of view of the state and local administrators, the primary requirement on management information system (MIS) data is comparability across schools and school districts, consistency over time and accuracy (conformity to the definitions and concepts used in that state's educational legislation and administrative regulations). When state legislation changes, the data that needs to be collected will change and consistency over time will be set aside in favor of accuracy.

Historical accident and different needs for data have resulted in MIS data that differ in concept and accuracy from state to state. Constructing valid estimates of national statistics through the aggregation of MIS data is, thus, no easy matter. Since estimates of national aggregates can be obtained where necessary by other means (e.g., sample surveys), the common core data program should not have the estimation of national aggregates as its central goal. The central goal of the common core data program should instead be to provide high quality data that are comparable across states. This implies that NCES needs to work cooperatively with the states to insure that data reported are consistent with the common definitions that have been adopted. If, however, specific data elements from some states are not comparable to the data from most other states, the data element should either not be published or be published in italics with a footnote describing reasons for the lack of comparability. When this occurs national aggregates should be reported only for the states that have provided comparable information.

III. THE NEED TO MEASURE EDUCATIONAL OUTCOMES

The statistics reported in The Digest of Educational Statistics and The Condition of Education are almost entirely measures of inputs into education rather than measures of the outcomes of education. Of the 70 tables in the 1984 edition of The Condition of Education, only 3 displayed data on test score outcomes of education and none displayed data on economic outcomes of education. A categorization of the tables in The Digest of Educational Statistics (see table 1) reveals that it did only slightly better in covering educational outcomes. Of the 173 tables in the 1983/84 edition, there were 9 tables displaying test score outcome data and 9 displaying data on economic outcomes (labor force status, occupation and earnings). The tests covered material that should be learned before entering college so all 9 of these tables can be considered outcome measures for elementary and secondary education. Of the 9 tables displaying data on economic outcomes, 3 tables addressed the experience of bachelor degree holders, 1 the experience of recipients of masters and Ph.D. degree holders, 3 the experience of high school graduates and dropouts and 2 tables presented data on all levels of educational attainment. There were no tables devoted to data on the labor market success of completers of 1 or 2 year occupational programs or associates degree holders. There were 3 tables on high school diplomas awarded and 4 tables on years of schooling completed. These, however, are measures of credentials awarded not of learning. The great variation in the educational standards across the country and the decline in standards over time call into question the usefulness of these statistics for comparisons across time and space of the performance of the educational system.[1]

Table 1 also presents data on changes over time in the subjects covered in The Digest of Education Statistics. The number of tables devoted to the economic outcomes of schooling has been stable over time. The number of tables providing test performance data rose from 2 in 1969 to 9 in 1983 due to the availability of data from the National Assessment of Educational Progress.

The Effects of Data Availability

There are important interactions between data availability and the evolution of public policy. Statistical agencies respond (often with a considerable

TABLE 1

COVERAGE OF THE DIGEST OF EDUCATIONAL STATISTICS

	1969		1975		1983	
	Nat.	State	Nat.	State	Nat.	State
Educational Outcomes						
Economic Outcomes	9	0	11	0	9	0
Test Score Outcomes	1	1	5	0	9	0
Educational Credentials and Inputs						
Credentials & Yrs of School Completed	3	5	5	3	4	3
Elementary Secondary Education Inputs	30	21	23	21	21	24
All Levels Inputs	11	2	12	3	6	2
Higher Education Degrees	4	2	10	2	19	1
Higher Education Inputs	36	13	35	11	31	14
Adult & Vocational Education	0	0	9	1	6	2
Federal Programs	5	2	6	3	3	4
International Education	6		6		8	
Libraries	7	1	8	2	5	2
R&D	6		7		0	
Equal Educational Opportunity	5		2		0	
	123	47	139	47	121	52
	170		186		173	

Odd numbered columns are counts of tables which present data for the nation as a whole. Even numbered columns are counts of tables which present data comparing states or in a few instances specific institutions of higher education or libraries. Input data includes information on students, staff, revenues, and expenditures.

lag) to the concerns of policy makers, but in addition, data availability has independent effects on the evolution of public policies.

The availability of statistics recording progress over time and providing geographical comparisons attract public and policy maker attention to a subject. The availability of data on inputs--numbers of students, numbers of teachers and expenditure--and the lack of data on outcomes contributed to an unfortunate tendency during the 60s and 70s for public attention and policy analysis to focus on a limited number of measureable inputs into learning rather on its outcomes.

The Nation at Risk report has made the quality and effectiveness of education the central educational policy question of the 1980s. The kind of data included in government publications is changing as a result. But it is doubtful that the National Commission on Excellence in Education would have had such an impact (and might not have even been appointed), if there had not been data available on (1) declining SAT scores and performance on NAEP exams and (2) the poor relative performance of American students in math and science in the International Education Agency's studies. The future of the educational reform movement will also be influenced by data availability. Since policy is made at the state and local level, there is a great need for performance measures that are specific to particular states and comparable across states and over time.

Test Score Outcome Data

The Department of Education has begun the process of supplying the need for comparative data on educational performance by making data on state means on the ACT and SAT scores and their changes over time widely available (Wall Chart and Indicators). The SAT and ACT tests, however, are curriculum free tests of aptitude for college and as such are not appropriate goals for instruction. They are also limited to college bound students. It is important that states be able to monitor their progress towards achieving instructional goals for all their students and across all grade levels.

Consequently, high priority should be given to publishing data on the following:

1. Performance on Achievement tests by subject, by age group (or grade), by minority status and by state. Many states mandate that all students in the state take particular achievement

> tests at specified points in their educational career. The problem is that different states use different tests. A study of the feasibility of linking together the various widely used achievement tests has been funded so a methodology for making the results comparable across most states will be available soon. Disaggregation by subject is important because it helps states monitor curriculum reforms that are specific to particular subjects.
>
> 2. Performance on competency tests necessary to receive a high school degree and how that has changed over time.

Consideration should also be given to publishing similar statistics for the larger urban school districts.

Colleges also need to be made accountable; therefore, data on the academic performance of their students need to be published. This could be accomplished by publishing Graduate Record Exam, SAT and ACT test scores for individual colleges and for all of the colleges in a state disaggregated by type of institution. Since the mean score will depend on the proportion of a college's graduates which take the Graduate Record Exam, it would be essential to publish this proportion along with the score.

Statistics like those in the wall chart titled State Education Statistics and those described above will inevitably be used to answer questions like "How are we doing relative to other states?" The public and elected officials will use such numbers to assess the effectiveness of the educational system in their state and to monitor the progress of state reform efforts. The publication of the wall chart suggests such uses are intended.

There are dangers, however. Some states are apparently planning to use the mean test scores for school districts as a very powerful tool of oversight.

> In the past year, four states--Arkansas, Kentucky, South Carolina, and Texas--have included provisions for academic bankruptcy in their reform bills. The bills in all four states ask that local school districts meet certain minimum levels of student achievement and other standards or face the possibility that the state will mandate additional reforms for the district, even including wresting control of the local school system from the local school board--the academic equivalent of putting the district into receivership. (Anderson and Pipho, 1984; p. 211)

If decisions of this magnitude were to be made without taking into account the demographic background of the students in the district, great injustices could be done. While state comparisons of performance could not conceivably lead to the federal government imposing changes on a particular state's educational system, they might have powerful political consequences. For example, they could influence the outcome of an election for State Superintendent of Public Instruction.

Comparisons of SATs, ACTs or achievement test scores across states probably reflect racial and social background differences of their population more than they reflect the quality of the schools in that state. Policy makers and the public need a statistic for comparing the performance of state educational systems that is less subject to such biases. Comparative data that purport to measure the performance of an educational system but in fact measure the talents and background of the students could confuse the public debate on education more than they enlighten it. Only a tiny minority of the users of such statistics will be sophisticated statisticians or educational researchers who will estimate multivariate models that control for the student background characteristics. Consequently, the data needs to be reported in ways which make the comparisons across states as fair and valid as possible. It is therefore essential that someone (NCES, NIE or a researcher under contract) be given the task of producing performance indicators that adjust for the social background of the students and the talents they bring with them at school entrance. The following should be considered:

3. Changes in the relative performance of particular cohorts of children from particular states. Such a statistic can be constructed by putting scores from tests administered in the first and later grades on a common metric (e.g., standard deviation units, grade equivalents or state rankings) and then examining how the states performance on this ranking changes as grade level increases. While such comparisons might be made for different tests given at a point in time, comparisons would be more valid if the cohort was held constant. This would be done by calculating state rankings on 1st grade scores in 1976 and then comparing them to the rankings on 8th grade scores in 1984. Grade equivalents and standard deviation units will produce different results. If grade equivalents are to be reported, standard deviation units should be reported as well and the difference between the two should be explained.

4. Achievement test scores (levels and gain scores) that have been adjusted for the demographic composition of the states students. Another way of reducing the bias problem discussed above is to estimate statistical models predicting achievement as a function of student background characteristics using state

aggregate data and then report only the residuals from such a model.

Assessing Higher Order Skills

Even though most basic skills should be learned prior to the end of junior high school, many students arrive in high school without a through grounding in basic skills. This fact and great public stress on the importance of basic skills seems to have resulted in high school teachers coming to view the teaching of basic skills as their number one goal. When 10,360 high school teachers surveyed as part of the HSB supplementary survey were asked to rank 8 different goals of their teaching, 44 percent placed "Basic literacy skills" at the top (see table 2). Only 19 percent gave first rank to "Academic excellence or mastery of the subject matter of the course." High school should be a place for developing higher order skills and reasoning ability such as inference, analysis, interpretation, and applying general principals to specific cases. The growth of testing and the tendency of most tests to focus on lower rather than higher order skills may be contributing to a narrowing of the high school's teaching agenda. The National Council of Teachers of Mathematics is concerned about this trend and has called for a revision of both the mathematics curriculum and the tests used to assess progress. It recommended that:

- "Problem solving be the focus of school mathematics in the 1980s
- Basic skills in mathematics be defined to encompass more than computational ability
- Stringent standards of both effectiveness and efficiency be applied to the teaching of mathematics
- The success of mathematics programs and student learning be evaluated by a wider range of measures than conventional testing" (National Council of Teachers of Mathematics, 1980; p. 1)

Similar problems exist in other subject areas.

Testing can be designed to assess higher order skills. The recent revision of The California Achievement Tests had better assessment of higher order skills as an important objective. The College Board and The National Assessment of Educational Progress have successfully implemented standardized methods of assessing writing ability. An effective way to insure that we do not forget

TABLE 2

THE GOALS OF SECONDARY SCHOOL TEACHERS

	Rank							
	1	2	3	4	5	6	7	8
Basic literacy skills (reading, math, writing, speaking)	44	18	12	8	7	6	4	2
Academic excellence, or mastery of the subject matter of the course	19	18	16	12	11	9	9	6
Citizenship (understanding institutions and public values)	4	7	9	12	16	21	22	9
Specific occupational skills	4	5	7	8	9	12	21	35
Good work habits and self-discipline	18	28	22	14	9	5	3	1
Personal growth and fulfillment (self-esteem, personal efficacy, self-knowledge	13	14	16	19	16	13	7	3
Human relations skills (cultural understanding, getting along with others)	6	10	13	17	20	18	12	4
Moral or religious values	9	6	6	7	9	11	17	35

NOTE: Cell entries are the percent of teachers who assigned the goal a given rank. Except for rounding error each row should sum to 100. Columns sometimes sum to more or less than 100 because some teachers gave different goals the same rank.

The question was: "If you had to choose from among the eight goals for students listed below, how would you rank them according to their importance in your teaching? Enter a "1" for the most important goal, a "2" for the next most important goal, and so on, through "8" for the least important goal. (FIRST, RANK EACH GOAL. THEN MARK THE MATCHING CIRCLE NEXT TO EACH RANKING.)"

about the need to teach higher order skills is to publish time series and state comparison data from tests and subtests that tap higher order skills.

International Comparisons of Educational Performance

New data are becoming available comparing the academic performance of students in different countries. NCES should make sure that the results of these studies are accessible to educational policymakers.[2] A start has been made by including data from the Second International Mathematics Study in Indicators of Education Status and Trends. The following additional tables of International comparison data should be included in one of NCES's publications:

- Mathematics
  - comparisons of performance disaggregated by whether the item taps computational skills or problem solving ability;
  - the performance of students from each country that participated in the 2nd International Mathematics Study;
  - the changes in math performance between the first and second mathematics studies in the U.S. and other countries disaggregated by subject and by whether the item taps computational skills or problem solving.
- Science--The Second International Science Study
  - comparisons separately for biology, chemistry, physics, earth science, etc. and separately by whether the item asks for a fact or the use of higher order skills;
  - comparisons of the changes between the 1st and 2nd International Science Studies disaggregated in the same way
- Reading Comprehension (Thorndike, 1973)
- Literature (Purses, 1973)
- Civic Education (Torney, Oppenheim, and Farnen, 1975)
- Civic Attitudes (Oppenheim and Torney, 1975)
- Comparisons of Grade 1 and Grade 5 students in the U.S., Japan and Taiwar (Stevenson, 1983).

There is also a need for the publication of data on the school input and processes and environmental variables that may be accounting for the differences between countries in achievement.

- Time Engaged in Study of Particular Subjects (over the course of a student's academic career)

- Time on Task (percent of time in the school that is actually spent engaged in learning) Stevenson (1983) reports observational data on classroom time on task and parent child interaction for U.S., Japan and Taiwan.

- Time Spent Watching TV

- Number of Books Read

- Other Measures of Educational Process and Outcomes

- Measures of the Economic Outcomes of Schooling

The role of NCES and NIE in international comparative studies should be to stimulate others (e.g., the appropriate NIE Centers) to undertake such studies and to provide grant support for researchers engaged in such research. Probably the most critical need at present is for the collection of internationally comparable time use data for both children (in and out of school), parents and other adults (Stevenson, 1983; Stafford and Duncan, 1980; and Juster and Stafford, forthcoming).

Assessing Work Habits and Self Discipline

High school teachers ranked "good work habits and self discipline" as their number two goal (see table 2). Sixty-eight percent of the teachers reported it ranked among the top three of their teaching goals. Only 53 percent placed academic excellence in the top three and 43 percent awarded top three status to personal growth and fulfillment. Amatai Etzioni has said

> Thus, the root problem is not that millions of high school graduates have great difficulties in reading, writing, and 'rithmetric; these all-too-common dificiencies are consequences of insufficient self-discipline, of inadequate ability to mobilize self and to commit. These graduates enter the adult world twice handicapped. They suffer both from continued psychic underdevelopment and from the inadequate cognitive preparation this underdevelopment helped to cause. (Etzioni, 1984; p. 18)

When 167 employers in the Los Angeles area were asked what was the most important factor in job success, 63 percent picked "good work habits," 23 percent picked "technical job skills," and only 14 percent picked "linguistic and computing ability" (Wilms, 1983).

Clearly it is important to keep track of how well our schools are developing good work habits and self discipline. High School and Beyond contains a number of questions about study habits, class cutting, having work done on time, tardiness, attitudes toward study and beliefs about one's ability to control one's fate (self efficacy) that can be used to assess character traits. NCES publications should present some of this data and where possible make comparisons with the Class of 1972 data. It is important that these items be retained in NELS88 and that there be no changes in wording.

Self Efficacy

The locus of control or personal efficacy scale measuring the belief that one's future is controlled by one's actions is probably the best documented and most researched of the sociopsychological scales contained in HSB. It probably has a stronger coorelation with what Etzioni is referring to as self discipline than any other widely used scale. The scale fits nicely into economic theory, for it can be interpreted as a measure of the perceived risk and uncertainty surrounding the consequences of one's actions. There is a well documented positive coorelation between self efficacy--believing that one controls one's own fate--and labor market success. The analysis of longitudinal data containing repeated measures of self efficacy and labor market outcomes has shown that an important part of this association is a consequence of self efficacy's impact on employment and wages rather than the reverse (Andrisani, 1978; Duncan and Morgan, 1979). Kang and Bishop's (1985) analysis of followup data on HSB seniors found that self efficacy measured in the senior year had a larger impact on wages and employment in the 2 years after graduation than test scores, grades and virtually all other school items except curriculum and employment during high school. A one standard deviation rise in self efficacy raised earnings of women by 8.1 percent and the earnings of men by 6.7 percent.

Not surprisingly self efficacy also has important effects on learning. Using a cross lagged design in which senior test scores were a function of sophomore self efficacy, deportment, educational and occupational plans, self esteem and performance on verbal, math, science and social science tests and a host of other background variables, Hotchkiss (1984) found that a one standard deviation improvement in efficacy increased the verbal and civics test score of seniors by 6 percent of a standard deviation and the science test by 5 percent of a standard deviation.

Hotchkiss (1984) has also demonstrated that schools have important impacts on the development of self efficacy. In analysis of HSB longitudinal data in which self efficacy as a senior was modeled as a function of self efficacy as a sophomore, a long list of individual characteristics and a vector of school dummies it was found that school dummies exercised significant effects. The multiple partial correlation indexing the impact of the school attended was two-thirds the size of analogous multiple partial indexing the combined effects of personal characteristics and socioeconomic background.

These results suggest that self efficacy deserves an important role in theories of learning and in theories of occupational choice and career attainment. The results cited above should be viewed only as preliminary forays into the development of such theories. Almost nothing is known about how schools and teachers foster a sense of self efficacy. NELS88 will be a much better data set studying these processes than HSB.[3] Consequently it is critical that the self efficacy scale remains in NELS88.

Measuring the Economic Outcomes of Schooling

Improved performance on tests of cognitive ability and greater achievement in core academic subjects are not the only goals of elementary and secondary education. Raising the employability of youth is another important goal. The National Commission on Excellence in Education justified its recommendations for reform in part by the following:

> Knowledge, learning, information, and skilled intelligence are the new raw materials of international commerce and are today spreading throughout the world as vigorously as miracle drugs, synthetic fertilizers, and blue jeans did earlier. If only to keep and improve on the slim competitive edge we still retain in world markets, we must dedicate ourselves to the reform of our educational system for the benefit of all--old and young alike, affluent and poor, majority and minority. Learning is the indispensable investment required for success in the "information age" we are entering. (p. 7)

If schooling influences the productivity of the work force as the National Commission claims, then people with greater amounts and higher quality education will be more likely to find good jobs that pay higher wage rates.

Research on the links between schooling and later success in the labor market have found that test scores have rather small impacts on employment and wage rates (Bishop, 1985) and that improved performance on standardized tests accounts for only a small part of schooling's total impact on labor market success (Olneck and Bills, 1982). Vocational education has work attitudes, job readiness and occupational skills as explicit goals and the time devoted to developing these abilities may reduce the time devoted to the traditional academic subjects. Consequently, data on the economic outcomes of elementary and secondary education are essential.

Experience with the Vocational Educational Data System suggests that the schools probably cannot be expected to collect and report data on the economic success of their graduates that is comparable across schools and across states. Data on economic outcomes must therefore be obtained from the Current Population Survey, the Census, longitudinal surveys like NLS, and organizations like the College Placement Council.

Survey of Salary Offers to Associate Degree Recipients

The College Placement Council collects data from the placement offices of colleges and universities on the offers received by B.A., M.A., and Ph.D. degree recipients. Northwestern University's Endicott survey publishes similar data obtained from a survey of employers. There is no comparable national survey of salary offers to the completers of two-year associate degree or one-year certificate programs. The Middle Atlantic Career Counseling Association (MACCA) has been collecting these data on salary offers received by recipients of the Associates Degree from about forty 2 year colleges in their six-state area for the last ten years. Their data cover about one-sixth of the nation. Data from the VED system is inadequate because the lag in reporting salaries is too long and because salary data was not obtained separately for training programs of different length.

After getting support from the American Associations of Community and Junior Colleges (AACJC), MACCA and other interested groups NCES should contract with the College Placement Council (or an organization like the National Center for Research in Vocational Education) to conduct a salary offer survey in as many 2 year institutions as are willing to participate. Placement officials at participating institutions would report on the job offers received by graduates they provide placement services to. The more rapidly such information is disseminated the more valuable it is so semiannual reports should be released. Since the labor markets for AA recipients are regional separate reports should be published for each of the 10 census regions. The cost of such a data collection effort would be under $100,000.

New Tables Based on Existing CPS Data: Time Series

The tables in the 1983/84 Digest are useful and should be retained. There is a need for new tables, however, that provide comparisons over many years of labor market success (employment, occupation, and earnings) for recent school leavers. The data proposed would provide a means of tracking whether the decline in SAT scores and other indicators of academic achievement have resulted in a parallel decline in the quality of the jobs obtained by recent high school graduates and a rise in their unemployment rates. The following tables are recommended.

1. Unemployment rates and employment to population ratios of young people who graduated from or dropped out of high school during the previous year and are not enrolled in school (see table C43 of Labor Force Statistics Derived from the CPS: A Data Bank, 1982). If possible separate scores for blacks and for single females should be published. The table should also contain an unemployment rate for prime age (25-55) males and for prime age females as a point of comparison.

2. Indicators of the quality of jobs obtained by recent high school graduates and dropouts. The following are recommended:

   - Industrial Composition
     - share of jobs in manufacturing, mining, construction, transportation and public utilities
     - share of jobs in government
     - share of jobs in wholesale retail or service
   - Occupational Composition
     - share of jobs in laborers or service occupations
     - share of jobs in operative occupations
     - share of jobs in clerical occupations
     - share of jobs in sales occupations

   These data should be presented both for those who graduated and those who dropped out of high school the previous year, and for all 18-24 year old high school graduates not enrolled in college and for all 18-24 year old high school drop outs.

3. Average yearly earnings of young people not enrolled in school categorized by educational attainment. Three age groups should

be reported: 18-24, 25-34 and 35-65. The educational attainment categories might be 0-11, 12, 13-15, 16 and 17+. Separate data should be provided by sex and for full time full year workers. In order to accentuate the comparisons across educational levels, the information could be presented as ratios to the earnings of high school graduates (real dollar amounts of earnings would be presented only for high school graduates). Tabulating these data for recent school leavers and comparing it to average weekly earnings of other groups will provide a measure of the relative labor market success of those who have recently completed their schooling.

4. Unemployment rates for people categorized by education and by age (e.g., 18-24 and 25-65).

Data on the labor market behavior of students is also important, though not as critical as data on recent school leavers. High school and college students spend a great deal of time working.[4] These jobs have both good and bad effects on educational outcomes. They help finance college but they do so by often reducing the time available for study.

Improvements in CPS Data

NCES can greatly improve the usefulness and timeliness of data from the Current Population Survey's (CPS) October supplement on attendance and completion of schooling by arranging for the addition of the following questions:[5]

- usual weekly earnings (same wording as in the January and May supplements)
- highest degree or credential received
- whether school from which highest degree received was a 4 year college, 2 year junior or community college, or technical institute
- whether the school most recently attended (or from which highest degree was received) was public, Catholic, other religiously oriented private or sectarian private
- the students major or concentration at that school (for high school the categories might be the same as those in question 2 of the HSB senior questionnaire, greater detail would be obtained from those with some college or a college degree)
- whether the individual received special education services in elementary or secondary school and the type of handicap

- state in which high school was located
- state in which college was located

The addition of these 8 questions would greatly improve our ability to track changes over time in the economic returns to various types of schooling and to compare the effectiveness of schooling in different parts of the country. This small addition to only one month of the CPS would go a long way toward meeting the Congressionally mandated requirements for data on the outcomes of vocational education.

The monetary returns to college attendance and to majoring in specific fields vary over time due to shifts in demand and supply for the skills taught in these schools (Freeman, 1971; 1971a; 1971b). The monetary benefits of completing high school and of receiving different types of high school preparation also vary over time. Estimates of the impact of vocational education of the earnings of 1980 high school graduates (HSB) are considerably more positive than estimates in earlier cohorts (such as Class of 1972, the NLS Young Mens survey, and the NLS Youth Cohort) (Kang and Bishop, 1985; Meyer, 1982; Daymont and Rumberger, 1982). The change could be a result of improvements in the quality of vocational offerings but it might also have been a consequence of the unique economic environment that greeted 1980's high school graduates. Data from repeated CPS surveys are necessary if we are to distinguish long term trends in the return to vocational or other particular types of education from short run shifts due to the business cycle. The CPS obtains earnings information on a much larger sample (over 70,000 households) than other surveys available for studying the impacts of curriculum and private or public control of institution (e.g., NELS and NLS) and simultaneous analysis of data from more than one year increases the sample size even further. This feature means that CPS data can provide reasonably reliable estimates of economic outcomes for smaller subgroups (e.g., male graduates of Catholic high schools who did not go on to college or female vocational education students in nontraditional occupations). The cost of such a supplement would be about $50,000 for data collection, another $150,000 for programming, analysis and writing a report.[6] For this rather modest investment we would have data on the economic benefits of different types of schooling for more than 100,000 working adults. The cost would be under 2 dollars per person.

It is recognized that the classification of individuals by their majors or field of study would not be as reliable as the data that could be obtained from transcripts. Other longitudinal data sets which merge followup data on economic outcomes with high school and college transcript information and baseline measures of aptitude, academic achievement, and career plans would still be needed. These data sets cost at least 200 dollars per observation, however. To be useful there must be data on comparison groups (i.e., those who do not take vocational courses) as well as on those who complete the program that is the object of interest. Longitudinal studies like NELS88 cannot be conducted on the scale or with the frequency that would be needed to construct a reliable time series, to make comparisons between the larger states or to provide estimates of economic outcomes for subsets of the population. Consequently, important as the longitudinal surveys are for analyzing the effects of various educational programs the augmented CPS would be a useful supplement.

Comparing Economic Outcomes in Different States

Indicators of economic outcomes for specific states are also needed.[7] Differing economic climates need to be taken into account, so tables must be structured so as to highlight statistics that compare the labor market success of groups with differing amounts or kinds of education but subject to the same economic climate. The greatest need is for tables which describe how particular increases in educational achievement (e.g., graduation from high school, completing some college, completing a four year degree) improves one's labor market success in that state. Since most people find employment in the state where they complete their education, these tables will provide a measure (an imperfect one to be sure) of the economic payoff to that level of schooling in that state--something that should be correlated both with the quality and the labor market relevance of that level of schooling in that state.[8] Since the source of data for this would be the 1980 Census, the publication of this data would not need to be repeated each year. The tables that should be developed and made more accessible to state policy makers are the following:

- Earnings of different age groups (18-19, 20-21, 22-24, 25-29, 30-34, 35-44, 45-54, 55-64, 65+) by years of schooling completed by sex, by minority status and for all workers and full time full year workers. (Available in Table 237 of Detailed Population Characteristics of the 1980 Census.) The focus of the table should be the earning ratios for people with differing educational attainment.

- Occupational and industrial distribution of people categorized by educational attainment, age, sex, and minority status.
- Unemployment rates and employment to population ratios by people categorized by educational attainment, age, sex, and minority status.

The usefulness of the 1990 Census for such comparisons would be greatly improved if questions were added on the field of study in high school and college, degrees received, state in which the individual attended high school, and state in which the individual attended college.

IV. NEED FOR DATA SETS FOR ESTIMATING CAUSAL MODELS OF EDUCATIONAL IMPROVEMENT

Probably the least controversial conclusion of effective schools research is that schools are institutions that have cultures (norms and expectations that students and teachers hold in common) which have powerful effects on how teachers teach and what students learn. Many of the most important outcomes of schooling--verbal ability, reasoning ability, self discipline, character, employability, ability to work as part of a team--are the result of the entire school experience and not something learned in particular courses. There is a need, therefore, for research that takes the school as a unit of analysis. Two very important issues need to be addressed.

- What is it about a school that causes students to learn more, to develop better character, or become more employable? Associations have been found between school climate and learning and lists of school effectiveness characteristics have been generated. The lists seem reasonable but the research that has generated them suffers from serious methodological weaknesses (Purkey and Smith, 1982).
- How are effective schools created? How does a school's ethos evolve? Since the school characteristics that are associated with school effectiveness are not under the direct control of a principal or school board, we need to describe the relationship between administrative actions and school climate.

Data for Studies of School Impacts on Learning, Character Development and Employability

The first issue can be addressed if a dataset has the following features

1. Longitudinal data on students while they are at school.
   - achievement test scores at various points in time
   - aptitude test scores
   - family background
   - educational and occupational plans
   - work habits, study habits, self discipline and self efficacy
   - courses taken from transcripts
   - career goals
2. Longitudinal data on students after they leave high school.
   - college attendance and completion
   - employment, wages, on-the-job training, productivity, job search and turnover

- crime and drug/alcohol abuse
- career goals, locus of control and self esteem
- job characteristics (e.g., cognitive complexity, autonomy, status)

3. Comprehensive description of the character of the school (inputs, processes, climate) while the student was at the school
   - interviews with principal
   - interviews of teachers and other staff
   - aggregated responses of students describing school climate

The National Educational Longitudinal Survey of 1988 described in the February 1985 NELS88 Position Papers (Longitudinal Studies Branch, NCES 1985) has the three features necessary to support studies of how schools effect the learning, character development and labor market success of high school students. Below is a list of comments on specific aspects of the proposed design.

- Substituting data collection on entering college students for data collection on graduating high school seniors is recommended, for it makes study of college climate and teaching feasible. A survey of college teachers should be a part of the NELS88.
- An attempt should be made to follow up all of the NELS88 students after they leave school not just a sample. If cost is a deterrent, less frequent followups (e.g., at 3 year intervals rather than 2 year intervals) would be acceptable. Another alternative would be to drop the followup of recent college graduates and use followups of HSB and NELS88 instead.
- Many of the most important effects of a quality education do not appear until many years later. Followup should consequently be continued for 20 years. Once the cohort has reached the age of 24 the frequency of followup might be shifted to once every 5 years.
- School records should be accessed to obtain
  - courses taken
  - grades in specific courses
  - test score data on all widely used achievement, aptitude and minimum competency tests taken by the student for as far back in the students educational career as possible.
  - receipt of special education service and nature of disability
  - absenteeism
  - tardiness
- The teachers who are surveyed need to be selected in a manner which allows researchers to generalize either to a definable subset of secondary school teachers or to the entire population.

- Teacher assessments of individual students who are in the HSB survey are quite valuable. The checklist the teachers complete should include an item on locus of control.

- Consideration should be given to developing a data set which matches students to particular teachers. In a subset of schools arrangements could be made to track which teachers each student has over the course of high school. In these schools all teachers would be surveyed and the size of the student sample would be increased.

- The principal and guidance counselor should be asked whether they respond to employer requests for transcript information under two different circumstances; the student has signed a release form, the student has not signed a release form. Nationwide Insurance of Columbus Ohio sent more than 1,200 requests for transcript information with student releases to local high schools in 1982 and received only 93 responses. Not responding to such requests when former students have signed a release is a violation of the Family Educational Rights and Privacy Act of 1974.

Sophomore Baseline Questionnaire

- The locus of control (self efficacy) scale should be retained.

- Some students spend little time on homework because they find it easy to accomplish quickly. Others spend little time because they do not finish the assignments. Therefore it is important to ask what proportion of his/her homework assignments are completed and handed in on time.

- Drug and alcohol abuse is a serious barrier to learning and questions on these matters should receive high priority. I do not agree that these questions are too sensitive to ask in the sophomore survey. The questionnaire currently asks about being in trouble with the law.

- The proposed addition of new questions on disciplinary problems is a good idea.

- The question about when was the last time you worked (BB021) is not worded well: A question asking what proportion of the school year the student has been working is suggested to replace it.

Employment Data in Followups (applies to NLS72 and HSB as well)

- Continuous time employment histories as in HSB followups are very valuable and should be retained. The number of different jobs should not be limited to 4, however.

- The question on how you found this job is very valuable and it should be retained. There is evidence that for vocational students' informal recommendations by teachers are more effective than school placement services in finding jobs for students and graduates (McKinney et al. 1980). A distinction between the two should be drawn by adding "Through a former teacher" to the list of alternatives. Finding a job through a friend or relative who works at the firm has an especially strong tendency to reduce turnover. The distinction can be drawn by using the following response categories "through a relative who worked at the firm, through a relative who did not work at the firm, through a friend who worked at the firm, and through a friend who did not work at the firm."

- The question on why you left the job should be retained, but involuntary terminations need to be disaggregated into:
  - fired
  - temporary or seasonal job
  - lay off not anticipated when hired.

- The question on training (51A-D and 52A-E of Second Followup of Seniors) should be asked for as many jobs as possible and especially for the first few jobs after leaving school. If a tradeoff is necessary, reduce the number of job satisfaction questions.

- The unemployment section is good and should be retained. It should include a question about receipt of unemployment insurance. A separate section should be developed to ask about the search process that led to the first job after leaving school the last time.

- Questions on tardiness, absenteeism and "goofing off" on the job that parallel similar questions about school behavior would be desirable.

- Wage rates do not fully capture the productivity impacts of a good education. Studies based on absolute measures of productivity of incumbents in particular jobs that pay the same wage to everyone with equal tenure find it has a coefficient of variation of 20 percent (Schmidt and Hunter, 1983). The respondent should be asked to make a self-assessment of their effort on the job and of their productivity. Question wording might be something like: "Relative to the other employees at the firm, how much harder or less hard did you work? How do you think your supervisor assessed your productivity relative to the other employees doing similar work?" If a tradeoff is necessary reduce the number questions of satisfaction and probabilies of taking particular kinds of jobs.

- Consideration should be given to obtaining permission from students to speak to their employer or supervisor and then asking the supervisor for an assessment of how the individual's productivity and capabilities compare with other employees.

Data for Studies of School Improvement

Studies of how schools evolve require data sets with repeated measures of school climate and school processes. Such data sets are very expensive to create, however, because the key climate and process variables cannot be measured without interviewing large samples of teachers and students at each school. A recent survey of 10,360 teachers which measured school climate and processes at 500 high schools cost more than $200,000 to conduct. Longitudinal data on school climate and process can be made available at no additional cost simply by returning to these 500 schools in the NELS88 and repeating some of the same questions in the survey of teachers.

Data for Studies of Learning in Elementary Schools and Junior High Schools

Longitudinal data sets that start when the child is very young and follow him/her all the way through school and into the labor market would make possible a number of very valuable studies:

- Long term effects of early childhood education and types of special education services
- Elementary school effectiveness
- Junior high school effectiveness
- Retention in high school through the 10th grade. Most dropouts leave school before the end of their sophomore year and so are not part of the NELS88 baseline.
- Evolution of career awareness and aspirations prior to 10th grade.

Consideration should therefore be given to, starting in 1988, a longitudinal survey of second graders and their parents in a sample of the elementary schools that feed into the high schools selected for NELS88. The parent child data would then be merged with school records on grades and aptitude and achievement test scores and surveys of the principals and teachers of the school attended. These students would then be resurveyed in 1992 and 1996 and serve as the sophomore cohort of NELS96.

Another way a longitudinal data set might be developed would be to undertake a followup of the 1st through 3rd grade children who participated in the

Sustaining Effects Study in 1975 through 1978. An additional requirement would be surveys of the principals and teachers in the elementary and secondary schools attended by these students and the development of a mechanism for matching each student to the teachers that he/she had. Even if no further augmentation of the Sustaining Effects data is planned there would be significant benefits to making a well documented version of the data set more widely available through NCES.

FOOTNOTES

1. Awarding a high school diploma to a student who cannot read or write, may make that individual better off but it depreciates the value of the diploma for everyone else and lowers the efficiency of the process which matches individuals to jobs.

2. The IEA data comparing the U.S. and other countries first became available in the early 70s. To my knowledge these data never appeared in an NCES publication until after the report of the National Commission on Excellence in Education. If these statistics had been appearing regularly in publications like The Condition of Education, the nation's attention might have been directed to the problem of declining standards much sooner than it was.

3. NELS88 is a better vehicle for studing the effects of school process and climate on student outcomes because the school process and climate variables will be measured when the students are sophomores rather than 2 years after graduation as in the case of HSB.

4. High school juniors and seniors average more than 10 hours a week in paid employment during the school year. Time engaged in a learning activity averaged under 20 hours (Bishop, 1985; p. 9).

5. An alternative might be to add the questions about school background to a month that already has the weekly earnings data such as May or January. The March Survey is already too crowded to be considered for this purpose.

6. The source of these estimates is an off the record conversation with a former Census Bureau staff member).

7. At present the Digest of Education Statistics does not contain such data (see table 1).

8. The CPS question on state in which one received one's high school education described in the previous section would allow the construction of similar tables using CPS data that would assign migrants to the state where they received their education rather than their state of residence.

REFERENCES

Andrisani, Paul J. "Internal-External Attitudes, Personal Initiative, and the Labor Market Experience of Black and White Men." The Journal of Human Resources 12 (3) 1978:308-325.

Bishop, John. Preparing Youth for Employment. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1985.

Comber, L. C., and Keeves, J. P. Science Education in Nineteen Countries: An Empirical Study. Stockholm: Almqvist and Wiksell, and New York: John Wiley, 1973.

Daymont, T. N., and Rumberger, R. W. "The Impact of High School Curriculum on the Earnings and Employability of Youth." In Job Training for Youth, edited by R. Taylor, H. Rosen, and F. Pratzner. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1982.

Duncan, G. J., and Morgan, J. N. "Sense of Efficacy and Subsequent Change in Earnings--A Replication." Journal of Human Resources 14 (4) 1979:649-657.

Etzioni, Amitai. "Self-Discipline, Schools, and The Business Community." Final report to the U.S. Chamber of Commerce Foundation, 1984.

Freeman, Richard B. "A Cobweb Model of the Supply and Starting Salary of New Engineers." Industrial and Labor Relations Review 29 (January 1976a):236-248.

________________. The Market for College-Trained Manpower: A Study in the Economics of Career Choice. Cambridge, MA: Harvard University Press, 1971.

________________. The Overeducated American. New York: Academic Press, 1976b.

Juster, F. Thomas, and Stafford, Frank. "Time Use Survey." In Time Goods and Well-Being, edited by F. Thomas and Frank Stafford. Ann Arbor: Institute for Social Research, University of Michigan. Forthcoming.

McKinney, Floyd, et al. Factors Relating to the Job Placement of Former Secondary Vocational Educational Students. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1982.

Meyer, R. "Job Training in the Schools." In Job Training for Youth, edited by R. Taylor, H. Rosen, and F. Pratzner. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1982.

National Council of Teachers of Mathematics. An Agenda for Action: Recommendations for School Mathematics of the 1980s. 1980.

Oppenheim, A. N., and Torney, Judith. The measurement of Children's Civic Attitudes in Different Nations. Stockholm: Almqvist and Wiksell and New York: John Wiley, 1974.

Olneck, M. R., and Bills, D. B. "What Makes Sammy Run?: An Empirical Assessment of the Bowles-Gintis Correspondence Theory." American Journal of Education 89 (1) November 1980:27-61.

Purkey, Stewart C., and Smith, Marshall S. "Effective Schools--A Review." Madison: Wisconsin Center for Education Research School of Education, University of Wisconsin at Madison, September 1982.

Purves, A. C. Literature Education in Ten Countries: an Empirical Study. Stockholm: Almqvist and Wiksell and New York: John wiley, 1973.

Stafford, Frank P., and Greg Duncan. "The Use of Time and Technology by Households in the United States. In Vol. 3 of Research in Labor Economics, edited by R. G. Ehrenberg. Greenwich, CT: JAI Press, 1980.

Thorndike, R. L. Reading Comprehension Education in Fifteen Countries: An Empirical Study. Stockholm: Almqvist and Wiksell and New York: John Wiley, 1973.

Torney, J. V.; Oppenheim, A. N.; and Farnen, R. F. Civic Education in Ten Countries: an Empirical Study. Stockholm: Almqvist and Wiksell International and New York; John Wiley, 1975.

Wilms, Wellford W. Technology, Job Skills, and Education: A Reassessment of the Links. Los Angeles: University of California, 1983.

MONITORING THE CONDITION OF EDUCATION

by Alphonse Buccino
The University of Georgia

Executive Summary

This paper responds to the NCES invitation to participate in the process of redesigning the elementary and secondary education data program. The purpose of NCES is seen in terms of monitoring the adequacy of the Nation's education enterprise in meeting individual and societal needs and expectations. To do this, a sense of mission for the enterprise is essential. This mission is proposed in terms of producing leaders, technical specialists, and informed citizens, and for addressing equity concerns and optimal use of talent.

A framework for monitoring the education enterprise is described consisting of five categories of variables. These categories are: outcomes, participation, resources, impact, and purposes. Each of these categories is elaborated and defined. An argument is made that outcomes is the key category. Additionally, a caveat is presented in that the framework comprises an input/output model. Professionals in education typically dislike such models and prefer more expressive ones. Nevertheless, the input/output model is the only model that is well-defined. Moreover, it is understood and preferred by policy and decision makers and the public.

Suggestions are then made for data sets within each of the categories. Especially notable data sets discussed are the following ones: student achievement data (NAEP); course enrollment data; data regarding teacher quality as well as quantity; and data regarding curriculum content (e.g. commonly used textbooks).

Additional consideration is suggested for information in such special areas as: international comparisons; informal, out-of-school education; and technology in education. It is strongly recommended that NCES develop guidelines to facilitate compatibility among data collected at the state level and by other institutions and agencies. Additionally, special attention is urged for the encouragement and support of interpretations of data from NCES and other sources.

The paper concludes with a discussion of why data comparison is especially important. A focus on outcomes is urged, with other variables tightly linked to outcomes. Comparisons should be among states, within states, and among nations.

MONITORING THE CONDITION OF EDUCATION

by Alphonse Buccino

The University of Georgia

Mission of the Education System

The elementary and secondary data program of the National Center for Education Statistics (NCES) should be designed to monitor the adequacy of the Nation's education enterprise in meeting individual and societal needs and expectations. To do this, a sense of mission for the enterprise is essential. One aspect of such a mission concerns the scope and level of knowledge and skill needed for the jobs of today and tomorrow. This refers to the adequacy of the education enterprise for maintaining the health and vitality of the U.S. economy. Two groups of particular interest in this regard are leaders and technical specialists. The significance of this aspect of the mission is reflected in such national reports as _A Nation At Risk_ (National Commission on Excellence in Education, 1983) and _Educating Americans for the 21st Century_ (National Science Board Commission on Precollege Education in Mathematics, Science and Technology, 1983).

But education can contribute to the national welfare through the preparation of leaders and technical specialists only to the degree that their work is understood by citizens and can be assimilated into our social structures. Accordingly, our education system must also produce informed citizens.

Additionally, education has long been the route by which upward mobility has been achieved by disadvantaged groups in our society. This has not changed. Thus, our education system must be instrumental in aiding those who have been excluded from full realization of their capabilities. This requirement for the education enterprise is stressed both for reasons of equity and to increase the size of pool of talent from which future leaders and technical specialists are drawn.

The inclusion of informed citizens and equity concerns in the foregoing is a reaffirmation of the Jeffersonian principle that an enlightened citizenry is the only safe repository of control over the ultimate processes of society. This principle is clearly articulated in the reports of the National Commission on Excellence and the National Science Board Commission that were referred to earlier. Today, the issue of the adequacy of the education enterprise in addressing these purposes refers to the goal of a fully educated citizenry, not just a long schooled one (Resnick and Resnick, 1985).

In summary, a program for monitoring the education system of the U.S. must address the adequacy of the system for producing leaders, technical specialists, and informed citizens and for addressing equity concerns and optimal use of talent. Included within the definition of adequacy are considerations regarding numbers of students, cost-effectiveness, and the general health and capability of the system itself.

A Monitoring Framework

There are many ways to describe the education system and assess its adequacy. To help focus in on an approach, it is useful to consider categories of variables. The discussion here is an adaptation of the framework set forth by Dennis Gooler (1975). This organizing framework for the development of data or data series consists of five categories.

The first category we present is outcomes. This has two broad components: tests and credentials. Tests provide measures of what people know and what intellectual and performance skills they possess. Use and misuse of tests is a matter of considerable controversy. Accordingly, the use of tests and test data requires a balanced and thoughtful scrutiny. This issue is further discussed later in the paper.

The other component in this category, credentials, concerns such things as diplomas, certificates, and degrees awarded by the education enterprise for successful completion of programs or curricula. Patterns of credential earning also provide measures of outcomes of education.

I begin with the outcome category because I believe it is the most important and should constitute the focus of any program to monitor the education system. Moreover, the categories below should be developed in such a way as to link tightly to outcomes.

A second category is participation. This category addresses the question: Who does education serve? It concerns how many and what kinds of people take part in education and the form it takes. Included here are such elements as retention rates of education programs and patterns of curricular interests of students. My experience is that this category is second in importance only to outcomes and is essential to an understanding of outcome information.

A third category is the resources available to education. This may include personnel acting as teachers, institutions offering courses of instruction, or tax dollars. Resources may take the form of capital, personnel, or teaching materials. It may also include such things as types of educational programs offered, curricula, and the amount of time spent on them.

Resources available to education may well include educational research and development in that these contribute to improving education. Thus, measurement of support for educational R & D or of information portraying the scope and condition of the educational R & D enterprise should be considered in this category.

A fourth category of the proposed organizing framework concerns impact. This addresses the longer range effects of education extending beyond outcomes as discussed earlier. Impact addresses such questions as: To what uses do individuals put their education? What happens to individuals who participate in education? There are three types of relationships one might consider. One is the relationships among education, work, and income. Another type concerns such things as concepts of self-worth, participation in community affairs, and life styles. A third type might concern the impact of education on general social, cultural, and economic affairs.

The fifth and final category might be labeled purposes. This category would include data or data sets that provide descriptions of the needs of the education enterprise and expectations held for it. From these, of course, a general sense of mission for education can be derived and a set of purposes education is expected to serve. The preceding section offered one vision of the mission of the U.S. education enterprise. This category is explicitly included because there is a tendency to take purposes for granted. While a great deal of effort need not be devoted to this category, it is nevertheless useful to include in order to assure common understanding.

A caveat is in order about the framework presented here. Each age and culture defines education in terms of the meanings it gives to teaching and learning, and these meanings arise in part from the metaphors used to characterize teachers and learners. In the ancient world, one of the defining technologies was the potter's wheel with the student's mind; the analog of clay in the hands of the teacher. Later the defining technology was the mechanical clock (Kilpatrick, 1985). Other metaphors have likened education to agriculture or young students' minds to blank slates or empty vessels.

The framework presented above quite clearly embodies an economic metaphor or input/output model. A pejorative description is the "factory metaphor." One should realize that educational professionals and practitioners generally dislike this kind of metaphor and prefer more expressive ones that they believe more accurately reflect the subtlety of education.

However, there is the problem of going from metaphor to categories. I am not aware of any effort that completes an alternative metaphor as well as Gooler has rounded out the economic model. Moreover, the model presented here is recognizable by the general public and by decision and policy makers outside of education and probably preferred by them. As indicated earlier, outcomes is the key category. This view expresses strong agreement with Cooke, Ginsburg, & Smith (1985). There has always been a great deal of information (relatively speaking) about the inputs to education, but outcome information has been scarce and inaccurate. One of the most significant features of the current wave of public concern about education is a shift in focus from inputs to outcomes.

Suggestions for Data Sets

Rather than try to provide a complete list of data sets that might be included in each of the categories of our framework, I will focus on just a few that either are particularly important or represent something unusual in terms of what NCES has done in the past.

Outcome information is the key to any effective system for monitoring the condition of education in the U.S. In this regard, the issues raised by Cooke, Ginsburg, and Smith (1985) are germane. There is considerable interest in state-by-state comparisons, but the current information base for making such comparisons is quite inadequate. Progress is being made to allow state-by-state comparisons for National Assessment of Educational Progress (NAEP) results. This is very much to be encouraged as is argued further below.

Another issue is whether current tests are measuring the higher-order skills that are needed to function in a modern, technological society. There are also concerns that school curricula have narrowed to fit the narrow focus of tests on mastery of elementary-level basics even in the higher grades. Thus, there is a challenge to improve the adequacy of outcome measures and to do so in a manner that does not improperly distort the purposes of education.

With respect to participation, the amount of schooling people receive and drop-out rates are important items of information. But in my experience, the most sought after indicator has been course enrollment data. Three excellent surveys of course enrollments were carried out by NCES in 1948, 1960, and 1972 but none since then. Some information on course enrollment behavior is part of High School and Beyond, but it is not comparable to the aforementioned survey data and is not reliable insofar as it is self-reported. A regular program (say, every five years) for collecting course enrollment data is urgently needed. NCES might consider doing so in a less exhaustive manner than was the case for the 1948, 1960, and 1972 surveys in order to control costs and reduce respondent burden.

With respect to resources, I have already observed that information portraying educational R & D would be useful. Additionally, some thought should be given to the question of quality of certain resources, especially teachers. I note with interest that the preprint of Indicators of Educational Status and Trends (1985) attempts to do this. Additionally, the importance of information regarding the quality of the teaching force is emphasized by the National Research Council report (1985).

A matter of special interest is the flexibility of supply of teachers. While there are many reports of unfilled teacher positions especially in certain fields such as mathematics and science, the supply of degree holders in mathematics and science in jobs other than teaching is high (NSF, 1980). A National Education Association study in 1983 also noted that a large number of persons not now in teaching jobs have completed the requirements for certification as public school teachers. Thus, information is required that goes beyond the numbers of persons in teacher preparation programs and considers the attractiveness of teaching as a career in terms of salary and working conditions.

Another kind of information that might be sought in the resources category is information regarding the substantive content of programs (NRC, 1985). An excellent source or surrogate would be textbook usage. That is, information about which textbooks are widely used would provide information about the content of school programs. Publishers accumulate such information, but they frequently regard it as proprietary.

Impact information can probably be developed with other agencies of government (e.g. Departments of Commerce and Labor) that address issues of economic vitality and personnel needs. Inclusion of information in this category is especially important now and for the foreseeable future because of fundamental changes obviously taking place in labor markets. The decline of traditional manufacturing jobs and the rise of the information society create a situation where the kinds of academic abilities schools are best at developing appears to be increasingly in demand (Resnick and Resnick, 1985). Moreover, this category clearly expands the notion of what information is relevant to monitoring elementary and secondary education in light of a sensitive understanding of the mission of the education system.

With respect to purposes education is expected to address, I have referred to the National Commission on Excellence and the National Science Board Commission as sources. High school graduation requirements would be another indicator of what people expect of education. Additionally, expectations of the public and of prospective employers might be surveyed directly.

In summary, NCES is encouraged to emphasize the following data sets:

- Outcome information, especially as provided by the National Assessment of Educational Progress.
- Participation as provided by course enrollment data.
- Information on teachers in terms of quantity and quality.
- Information on curriculum content.

Additional Considerations

Some considerations in addition to the foregoing on information needed to address issues of monitoring elementary and secondary education are presented. The additional considerations are five in number.

International Comparisons. Current concerns about American education arise substantially from the belief that U.S. education is weakening, but other nations are improving their education. These nations include our most powerful economic (Japan) and military (U.S.S.R.) competitors. They have made strong national commitments to and have displayed considerable success in educating large proportions of their secondary school populations to higher levels of skill and understanding than is the case in the U.S. (NSF, 1980). This belief is expressed by both the National Commission on Excellence and the National Science Board Commission.

Nevertheless, there is some controversy about the belief as stated above. For example, Americans have greater access to higher education, and the American system is much more flexible in allowing individuals to move in and out of the education system over their entire lives. Thus, international comparisons between elementary and secondary education systems and practices may be inadequate. One must look at life-long patterns of education and the relation of education to individual and societal well-being. In any case, a data program designed to address emerging or continuing issues should include consideration of the international scene in some realistic way.

Informal, Out-of-School Education. Learning activities in schools are designed to be structured and intentional. Yet, there is increasing recognition that much that effects the quality of formal education occurs outside the classroom and beyond the control of the school. A great deal of learning takes place unintentionally and unconsciously through casual reading and other experiences, especially through television and museums (National Science Board Commission, 1983).

Accordingly, information regarding education in out-of-school settings should be studied. We have already mentioned television and museums as focal points. But libraries and community groups such as Boy Scouts and Girl Scouts should also be taken into account.

Technology. Technology, if we include such things as books and chalkboards, has always been intrinsic to education. But the computer and other complex technologies that have recently emerged are having especially significant impact on both the content and delivery of education (Buccino, 1985). It is desirable, and may even be necessary, to identify and collect explicit information regarding technology. This could include information on misuse or ineffective use of technology to the extent that these occur.

Data Compatibility. A great deal of data is collected at the state level and by other institutions and agencies. Unfortunately, little consideration is given to compatibility issues. This is so because the state-level data collectors have little guidance and are not in a good position to know how to design their own data collection for such compatibility. NCES can make a valuable contribution by developing explicit guidelines for state-level data collection that provide for making these data compatible among themselves and with the national data sets. Of course, following the guidelines would be voluntary.

Interpretation. Data cry for interpretation. When educational data are published in the press or reported on television, we are deluged with requests from the public to explain what they mean. Moreover, the publication of data, as such, can even be misleading in the absence of appropriate interpretation. Accordingly, I strongly urge that NCES engage other programs in the Department of Education and other federal agencies to encourage and support interpretations of data from NCES and other sources.

An interesting example regarding the need for careful interpretation concerns the impact of high technology on the knowledge and skill levels needed by workers. Some argue that as technological complexity of the U.S. economy increases, jobs at all levels increasingly require higher levels of knowledge and skill. But Levin and Romberger (1983) argue that proliferation of high technology industries and their products is far more likely to reduce the skill requirements of jobs in the U.S. than upgrade them. While questioning the impact of technology on the knowledge and skills needed for the economy, Levin and Romberger do argue for higher levels of education on the basis of the ideal of informed citizens and personal well-being.

To a certain extent, the Condition of Education and other NCES data publications have tried to do this. What I suggest here is something like a series of about five substantial interpretive papers annually to be published along with data sets. Such papers might well focus on emerging issues with a separate discourse for tracking continuing issues. The image here is a modified Condition of Education. It would have two parts. One part would present data in graph and table form organized in accord with the framework and categories suggested earlier. The other part would comprise a collection of about five substantial analytic and interpretive papers focusing on emerging issues and a review of status regarding continuing issues.

The Problem of Standards and the Importance of Comparisons

The paper is closed with a discussion of why data comparison is especially important. While it is relatively easy to develop criteria to monitor the adequacy or performance of the education system, it is much more difficult to develop standards -- which determine the degree to which criteria are met. A criterion is a characterizing mark or quality. In this context criteria might well be the categories forming a framework for assessing the condition and progress of the American education system that were discussed earlier. These are: outcomes, participation, resources, impact, and purposes.

A standard is an exemplar of a criterion, a definite level or degree of the quality, defined by the criterion, that is adequate for a specific purpose. For example, large may be a criterion, while for different purposes the Empire State Building, or a whale, or a meter may constitute the definite level or degree of largeness that is adequate.

Given a criterion, there are two essential methods for establishing a standard. One is to identify an objective exemplar. We do this, for example in the case of length, when we select a certain specific platinum-iridium bar to represent the length of one meter and place it in the Standards vault in Paris. Thereafter, an object is said to have length one meter if it is equal in length to the exemplar, the platinum-iridium bar in the vault.

But we do not always have an objective exemplar identified in this way to place in a vault. This gives rise to the second method for establishing a standard. This is the method of comparison. It is not always easy, for example, to determine whether a given object is large. However, it is usually quite easy to decide, given two objects, whether one is larger than the other.

Thus, generally speaking, it is usually quite difficult to determine the degree to which a given object possesses a specific quality. On the other hand, given two or more objects, it is relatively easy to compare them to each other with respect to the quality in question and assert the relative degree of the quality reflected in the objects.

This illustrates the fundamental problem confronted by NCES. Given such criteria as we have discussed, there are very few objective exemplars for measuring the degree to which a school system possesses these qualities. The only recourse is comparison, and comparison is controversial from a political point of view. If the condition of one school system is "better" than another, then -- logically -- the other's condition has to be "worse" than that of the first. It may be all right to be better, but it is usually unacceptable to be worse.

In this paper, we set forth a set of criteria (qualities) that NCES should use in monitoring the U.S. education system. Typically, exemplars for these criteria are not available. Accordingly, NCES must consciously develop a comparative approach in order to set forth standards to accompany the criteria. As we have suggested, the comparisons may be among states. They may also be among nations and within states.

References

Buccino, Alphonse et al. Curriculum and technology. A presentation to the IBM Seminar for Deans of Education at Syracuse, June 1985. University of Georgia, 1985.

Cooke, Charles, Ginsburg, Alan, and Smith, Marshall. The sorry state of education statistics. U. S. Department of Education, 1985.

Gooler, Dennis. The development and use of educational indicators. In Educational Indicators: Monitoring the state of education. (Proceedings of the 1975 ETS Invitational Conference) Princeton: Educational Testing Service, 1975.

Kilpatrick, Jeremy. Reflection and recursion. Educational Studies in Mathematics, 16, pp. 1-26, 1985.

Levin, Henry and Romberger, Russell. The educational implications of high technology. NIE Research Report, Institute for Research on Educational Finance and Governance at Stanford University, 1983.

National Commission on Excellence in Education. A nation at risk. U. S. Department of Education, Washington, D.C., 1983.

National Education Association. Teacher supply and demand in public schools, 1981-82. NEA Research Memorandum, Washington, D.C., 1983.

National Research Council. Preliminary report of the committee on indicators of precollege science and mathematics education. National Academy of Sciences, Washington, D.C., 1985.

National Science Board Commission on Precollege Education in Mathematics, Science and Technology. Educating Americans for the 21st century. National Science Foundation, Washington, D.C., 1983.

National Science Foundation and Department of Education. Science & engineering education for the 1980s and beyond. National Science Foundation, Washington, D.C., 1980.

Resnick, Daniel and Resnick, Lauren. Standards, curriculum, and performance: a historical and comparative perspective. Educational Researcher, 14 (4), pp. 5-20, 1985.

DATA NEEDS FOR SCHOOL POLICY IN THE NEXT DECADE

James S. Coleman
The University of Chicago

Prologue

Because there will likely be extensive redundancy in the recommendations made to NCES by various educational researchers, policy makers and interested parties, and because much of what I would write were I to be comprehensive would merely increase the redundancy, I have chosen not to be comprehensive. Instead, in Part I of the paper, I have focussed on a single kind of problem, one for which I believe there will be little redundancy with other recommendations. I do so not primarily to increase the distinct information my paper will transmit, but primarily to focus the attention of NCES on a very important set of prospective problems in education, and on the importance of a body of data relevant to those problems. Because of the importance of the problem to be discussed in this paper, I believe these measures will show strong effects on school outcomes. In addition, the measures focus on matters which are directly subject to policy intervention.

After Part I, which addresses this single problem, I will address in a Part II some additional points, only loosely related to the first and to one another, concerning NCES data collection activities. Because these points are somewhat disparate, I have separated them off into a Part II.

Part I: The School, The Family, and the Community

Part I of this paper is based on a single premise: that in the decades to come, elementary and secondary schools (and pre-elementary schools as well) will be unable to function successfully unless they regard their task as something beyond that of educating the individual student. More specifically, the premise is that unless the school comes to provide certain functions that have been traditionally regarded as provided by the home and the community, it will be increasingly unsuccessful in its task. This premise is not based on a notion that schools should take on additional tasks such as "the teaching of values", or other tasks, but rather that in order for schools to succeed at their central task of educating children, they must approach this task quite differently than they have in the past.

The form this part of the paper will take is to first provide justification of this premise, second, to discuss some of its implications for the functioning of the school, and finally to indicate implications for the kind of data that will be necessary to account for the degree of success of a school in its task, and to guide school policy.

Changes in society, and how they affect the school's task

My premise is that the success of schools in decades to come will depend on their being able to provide functions that have traditionally been the province of the community and the home. The premise is grounded in certain large-scale social changes that have taken place and are continuing to take place. It is these social changes which can defeat the goals of the school if the school continues to address these goals in the way schools have traditionally done. The proposition stated in its most general form is that in the presence of a changed social structure outside the school walls, the school itself must change if it is to accomplish the same goals it has pursued prior to the societal change. Stated in this way, the proposition is almost trivially true. It is the specifics which give informational content, and it is to those that I now turn.

I begin with the observation that schools have always been most successful with children from strong families. That has generally meant families from higher socio-economic status, families with a stronger educational background, and families in which the parents themselves provide a verbally rich environment. Consistent with this is the fact that younger children in a family achieve slightly less highly than does the oldest sibling, for it is the oldest sibling whose verbal environment as a young child has had the highest fraction of adults in its composition.

Schools have, however, not always been successful with children from well-educated and high socio-economic status families. Children from families disrupted by divorce do not do as well in school as children from intact families, and children from high socio-economic families in which the parents are inattentive or disorganized have traditionally been the "problem children" of elite boarding schools. At the other extreme, schools have often been successful with children from strong families in which the parents' education is limited. Schools were successful with many children of earlier generations from rural or immigrant backgrounds in which there was little education, but a high degree of interest in the children's education and a high degree of resolve to see children do well in school. Schools of today are successful for many children from poor families with limited parental education, when these families are strong and attentive to their children's success in school.

This leads to the second observation, that schools are more successful with children from strong communities than with children from disorganized or weak communities. The prototype of a strong community is the rural communities of a few generations past and some (though not all) of the ethnic immigrant neighborhoods of a few generations past. In those communities, the social norms reinforcing school goals supplemented the family's own resources, and aided the success of children whose families might otherwise not have had sufficient resources to insure their child's success in school.

The prototype of the disorganized community is the ghetto of modern central cities, in which illicit and illegal activities distract children from the goals of the school as well as those of their families. In such communities, the social norms conflict with school goals, run counter to the family's aims and undermine the success of the children whose families might otherwise have had sufficient resources to insure their children's success in school.

But it is not only ghetto communities which can undermine children's success in school. Any community with a high proportion of disorganized families, or with parents whose attention is so fully directed to their own problems that they give little time to their children, generates norms destructive to children's success in school. A recent semi-autobiographical novel of a 20-year old young man who grew up in Beverly Hills (title: Less than Zero) is instructive. His description of the youth culture in that community of high income, high-status, well-educated families is a description of drugs, sex, violence, and self-destructive narcissism.

A third observation is that families in American society are becoming less strong with each generation, less able to provide their children with the kind of resources at home that their parents provided for them. Indicators of this are many: high and increasing divorce rates, which show no signs of declining to earlier levels, the replacement of family-wide leisure activities by adult social activities, "children's activities," and youth culture activities for which age-specific music both plays an important part and indicates the separateness. The increasing fraction of mothers of pre-school children in the labor force reduces the transmission of parental cognitive resources to children. A general shift of attention to the mass media of entertainment by family members of all ages helps undermine family values and attracts attention both of parents and children away from those intra-family activities that have traditionally aided the family and in doing so aided the school.

Finally, a fourth observation is that some of these same social changes, together with others, have greatly weakened the local adult community served by a school, and have largely destroyed those norms, and the sanctions accompanying them, upon which families and schools have in the past depended. In addition to the social changes described above that lead to family disorganization and parental inability to reinforce the school's goals, there is the major social change in which fathers, and increasingly mothers, work outside the local community where their child attends school. This change, which takes parents out of the local community, removes the possibility of a strong set of community norms which can reinforce the school's goals. A complementary change has added to this effect, for in many places the schools too have moved out of the community. This has occurred in some places through school consolidation, in some places through school desegregation, and in some places through staff professionalization, which has moved teachers away from personal involvement in the community where they teach.

These four observations taken together point to the new challenges that have come to confront elementary and secondary schools, and will increasingly confront them in the future. Together, the first and third observations imply that schools will increasingly be populated by children from homes in which the resources that schools have depended on will be absent or not used in the service of school goals, and that schools pursuing their task as they have in the past will be less and less successful. Together, the second and fourth observations imply that the community surrounding a school will be decreasingly a support to the school's goals, increasingly an impediment, and that schools failing to modify their activities will find themselves with an increasingly unmanageable student population.

Implications for the successful functioning of a school

The changes I have described above leave children with less adult attention, less adult interest, and less adult control than has been true in the past. Schools which do nothing new will find themselves with children more psychologically isolated as well as with children more controlled by peers, commercial entertainment, and exploitative adults than in the past. To prevent this, I see two possible avenues for schools to pursue. One is to help strengthen and rebuild the social structures in home and community which have in part abandoned children to the school and to peers, and the other is to build compensating social structures as part of school activities.

The first of these strategies implies two tasks, one focussing on the home and the second focussing on the community. Stated quite generally, the first task is to involve parents in their children's education, a task which will result in greater expenditure of parental resources (such as attention and interest) on the child and the child's education than would otherwise be true. To carry out this task requires a shift in a direction opposite to that which schools and teachers have taken in recent years. It requires encouraging parents to become involved in the school, even at the cost of having to take parents' interest and demands into account. It requires removing the shield that many teachers and many schools have used to keep out parental interference.

The matter may be put in terms of two diagrams, which, at the risk of oversimplifying can serve as a useful mnemonic. Figure 1A expresses the current form of relation between home, child and school in most communities, with the child as the only link, while Figure 1B expresses the form of relation that is necessary if the school is to successfully involve parents in their children's education and strengthen the home's capability of reinforcing school goals.

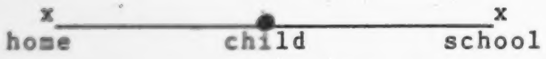


Figure 1A: School-Home Separation

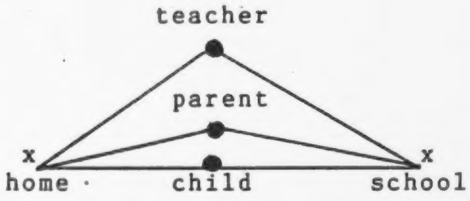


Figure 1B: School-Home Closure

Figure 1A is meant to denote a school-home relation which is entirely mediated by the child: The child has a relationship with parent within the context of the home, and with teachers within the context of the school, but these relations are separate and distinct. There is no linkage between school and home other than the child.

Figure 1B represents a situation I will describe as "school-home closure." It is meant to denote a school-home relation which is mediated not only by the child, but by either or both of two others: by the teacher, moving from the school into the home context, through home visits; and by the parent, moving from the home into the school context. For some parents who can or will spare little time for their children's schooling, this may mean only school visits or participation in school events. For others, it can mean involvement through volunteer services at the school.

Whatever the form of school-home closure, for the school of the future which pursues this strategy of strengthening the home environment, any of these activities should be accompanied by pedagogical activities from the school to the parent: Recognizing that many parents are cut off from those kinship and neighbor resources which can transmit information about the kinds of rules, practices, and facilities in the home that will help the child be successful in school, the school itself takes on the task of transmitting this information and encouraging the parent to use it.

The second task for a school which chooses the strategy of strengthening and rebuilding the home and community structures is a task which focusses on community structure. The aim in strengthening the community structure is to facilitate the creation of a set of norms and accompanying sanctioning mechanisms in the community that will reward those activities of children and youth which are in the direction of goals of schools (and parents), and negatively sanction those activities of children and youth which go against goals of schools. This is an orientation that was more prevalent in communities when the school-and-neighborhood more often constituted a functional community than is true today. In the fractionated communities that are found in much of America today, the absence of school involvement with the community is merely one indicator of the general decline of the community. This is not to say, of course, that a "community school" orientation is not to be found in some American schools, for of course it is. The present point is that this orientation can be especially important in strengthening those communities that are unable to reinforce school goals.

At the risk of oversimplifying, the kind of community structure which is able to develop norms and apply sanctions reinforcing schools' and parents' goals for children, a contrast may be made between community structures with what I will call "intergenerational closure" and those without such closure.* Structures that exhibit this closure can be described as those in which friends and associates of a child's parents are also parents of the child's friends and associates. As in the case of school-home closure, intergenerational closure may be shown by comparing two diagrams. In the diagrams, the vertical lines represent parent-child links across generations, while the horizontal lines represent friendship and associational links within generation.

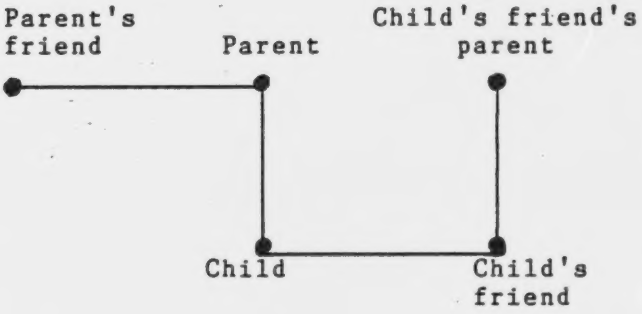


Figure 2A:
Intergenerational Separation

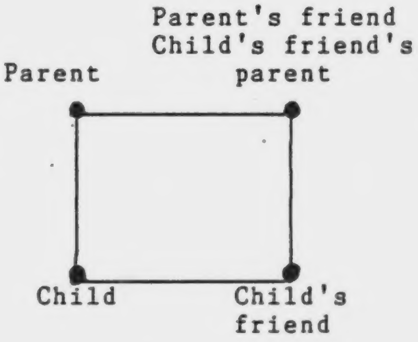


Figure 2B:
Intergenerational Closure

---

*In "Schools and the Communities they Serve" (Phi Delta Kappan, April 1985), I examine this contrast in greater detail.

Figure 1A characterizes a social structure in which there is separation between the community of children, focussed around the school, and the networks of relations in which parents find their friends and associates. In communities where most families have both parents working outside the community, and others are single-parent families, many of the friendship and association relations will go outside the local community, and the social structure will approximate Figure 1A.

Figure 1B characterizes a social structure in which networks of relations which connect adults are largely coterminous with the community of children in the school. The intergenerational closure that results makes possible a flow of information among parents about children, and about school. This flow of information, in turn, facilitates the growth of norms and the application of sanctions by the community, both positive and negative, which shape children's behavior. Parents will discuss what is acceptable behavior and what is not, parents and children will be congratulated for achievements of the child, and parents are not afraid to sanction the behavior of children who are not their own.

If a school chooses the strategy of strengthening and rebuilding the social structure of the community in a way that reinforces school goals, it will do so through attempting to create structures like Figure 1B, where the structure is currently like Figure 1A. In short, it must create and strengthen relations among parents of children in the school if those relations are to sustain norms that strengthen the school's goals. There are a variety of ways this can occur. The most obvious are parent's associations and PTAs, organizations which many schools do little to foster except where they arise naturally (which is principally in communities with structures like Figure 1B, where they are least needed). In private schools and public schools which are attended by choice, parents are sometimes required to commit themselves to some school event or activity which involves working with other parents. In various schools, there are parent-sponsored activities such as auctions and bake sales. In addition, however, ad hoc parent's groups are sometimes formed at a time of crisis around some problem area, such as drug or alcohol abuse.

A second strategy for schools confronted with weak or disorganized families, or with weak or disorganized communities, or with both, is to build a compensating social structure through and around school activities themselves. This strategy can be found most fully pursued in boarding schools, many of whose children are present precisely because of family disorganization or parental desire to be freed from daily attention to children's schooling. The social order established in these boarding schools may range from the hierarchial form of an Eton to the communitarian and egalitarian form of a Summerhill or an Ecole d'Humanite. But whatever the form, it is a social order, with norms, demands, and sanctions which surround its members, and which a school with a mimimum of sociological skill can shape. It is not a child-destructive social disorganization that exists in the larger society in the absence of strong families and strong communities.

Absent a boarding-school setting, schools beyond the lower elementary level may -- and many do -- attempt to capture the interest and involvement of children and youth through extra-curricular activities of various kinds. Many students whose unexceptional academic potential and lack of parental attention provides little incentive for intense involvement in school nevertheless do come to be intensively involved through some form of extra-curricular activity. In the presence of weakening community and family organization, some new pattern of extra-curricular activity may evolve in schools to bring a broader range of activities and interests, for a larger fraction of students, under the umbrella of school supervision.

In this section, I have described the implications for school functioning of the changing structure of the family and community for school functioning. In the next section, I will indicate some implications for data collection activities of NCES.

Implication for NCES data collection activities

The scenarios described above have various kinds of implications for NCES data collection activities. First are implications for new measurements that assess the kinds of social structural setting -- the kinds of family organization, the kinds of community organization, and the link between school and home and between school and community. Second are implications for measurement of school practices that act either to strengthen or to substitute for weakened home and community organization. Third are implications for ways in which NCES data activities themselves might augment parental resources, strengthening their capacity to aid their children's education.

Measuring the social structure and its relation to school: If the overall premise of this paper is true, data-gathering activities designed to provide information for school policy (like, for example, NCESs High School and Beyond) should obtain data that measures family characteristics, the school's relation to the family, community organization, and the school's relation to the community. Reasonably good measures of the first of these (which show strong relations to student performance) already are used in some NCES data-collection (e.g., High School and Beyond); it is the othe three that are largely missing.

The kind of data necessary can best be described by reference to Figures 1 and 2, for what is needed are measures of the relations that differentiate Figure 1B from 1A, and 2B from 2A. None of these are characteristically measured in NCES data-collection, though in High School and Beyond, there is one measure that can aid in distinguishing 1B from 1A, both at the individual student level and at the school level. This measure, obtained only incidentally and so far not used in analysis of HS&B data, is a question in the teacher comment checklist asking for each student in the sample (after a question as to whether the teacher knows that student) whether the teacher knows the parent. This allows measurement of the degree to which there is some form of school-home closure, though it does not allow distinguishing whether the closure occurs through the teacher in the home context or through the parent in the school context (see Figure 1B).

Additional measures which would obtain information directly relevant to school-home closure and intergenerational closure have been absent from NCES data-collection, though they could easily be included in instruments of the sort already used. Information on both types of closure could be obtained in student questionnaires of the sort used in HS&B and the National Longitudinal Longitudinal Survey of 1972 High School Seniors. When there are in addition parents' questionnaires or interviews (as in a subsample in HS&B), then even more direct and reliable measures relevant to the structures shown in Figure 1 and 2 can be obtained. (It is surprising, in fact, that in the HS&B parents' questionnaire, neither information on the parent's involvement with the school nor information on the parent's involvement with parents of other children in the school were obtained.)

I will not go into the particularities of just what kinds of instruments and items may be most useful for obtaining the relevant data, for that is relatively straightforward. The essential point is the recognition of what kinds of data are at issue here, and the potential importance of such data for assessing the functioning of schools in the coming decades.

Measuring school policies and practices relevant to social structure: In the earlier section on implications for the successful functioning of a school, I have indicated some of the kinds of school policies and practices that schools have initiated, and others that can be initiated, to alleviate the harmful impact of changes in social structure. This is only a beginning. Exploratory ethnographic studies and pilot studies can be initiated to discover the full panoply of such policies and practices that exist in American schools. Once such information is at hand, it can provide the basis for instruments or items that can measure the extent of these policies and practices. What is essential now is, as in the measurements described in the previous section, that the kind of data under discussion is clear, and that the potential importance of such data for explaining the differential success of schools is clear.

Use of NCES data activities to augment parental resources: In both the preceding sections, measurements were described which would have analytical value for policy-relevant research on school functioning. But there is another kind of value that NCES activities can have for the problems I have described. This is the encouragement and facilitation of parental and community use of information about student performance and school functioning.

Schools and school systems have been quite variable in both willingness to provide parents and the community they serve with data about student and school functioning, and their ability to provide such data. It was only through pressure from newspaper reporters that big-city school systems began to make public standardized achievement data at the school level. It was only Federal freedom-of-information legislation that gave parents rights to access to school records on their own children, and many schools discourage the use of these rights by parents. Yet this kind of discouragement is, if the premise on which this paper is based is correct, increasingly inimical to the successful functioning of the school. Parental resources, and interest in using these resources to benefit their child's education, can be amplified by free and easy access to information both about their children's progress and about the school's functioning. Community organization is more likely to be applied toward the improvement of education if facts which many school systems attempt to keep hidden (such as frequencies of various forms of violence, delinquency, and crime in the school, or the frequency of cutting classes or teachers' absence rates) were made public. An important role of NCES is to act, in effect, as a representative of the consumers of education with respect to information relevant to their interest. (In the past, NCES data services have been more use to education producers than consumers.) Some specific steps which can be of aid in this task are:

1. Publication of a booklet informing parents of their information rights vis a vis their children's schools, public and private, and giving information about how to interpret the usual items of information in school records. Such a booklet should indicate also information about school functioning that schools are required by state law to keep, or would normally keep as part of school management, with an indication of what kinds of information would, if parent groups can induce schools to make it public, be most valuable as indicators of school functioning (e.g., monthly teacher and student absence rates, yearly standardized achievement gains, dropout and transfer rates at each grade level).

2. Publication of material disseminated to school systems giving specifications for an appropriate system of provision of

consumer information, so that school districts that are so inclined will have a standard to turn to. Such a system can be designed to make use both of information required by the state department of education and of NAEP or NAEP-like information.

3. Design of a system of consumer information to accompany newly-introduced plans of school choice that states or school districts elect to introduce, either within the public school system or including non-public schools as well.

These are specific examples of the type of information services that NCES can provide and can stimulate which will have a direct effect in strengthening, not the family structure or community structure, but the ability of families and communities to support and aid their children's education. Schools have erroneously equated their comfortable insulation from parental and community pressures, and from the exercise of parental choice, with benefits to students. The educational establishment, NCES included, has done little to counter that self-serving action. But as the principal Federal information agency on education matters, NCES has both a responsibility and an opportunity to serve and protect the interests of consumers of educational services. This is the spirit of the large state-by-state comparison chart of educational inputs and outputs which NCES has just published. That spirit should be present in a much broader set of services, such as those described above, which can stimulate, encourage, and generate pressure for the opening up of information about school functioning to parents and community. Such services were less important when schools were closer to their communities, and when there were strong parental communities coterminous with communities of children and youth. But they are important now, and will become increasingly so in the future.

Part II

Included here are a few additional points about NCES data-collection and statistical activities.

1. The two major longitudinal studies of high school students, the National Longitudinal Study of High School Seniors of 1972, and High School and Beyond, covering high school seniors and sophomores in 1980, have proved to be extaordinarily fruitful longituidinal data bases - and have shown the value of the general design used by NCES in 1972 and improved in 1980. To continue to monitor the functioning of American high schools by continuing this series with subsequent cohorts is very important.

In addition, the experience gained from NLS-72 and HS&B, and the general value which these data bases have shown, should encourage the initiation of comparable series at lower grades of school. In general, it appears quite useful to concentrate NCES resources on obtaining and maintaining longitudinal data bases on comparable cohorts at periodic intervals, as in the case of NLS-72 and HS&B.

2. As part of the design of HS&B, an approach called "pluralistic policy research design" was used in modification of instrument and study design. (See "Policy Issues and Research Design," Report to NCES October 1979, by James Coleman, Virginia Bartot, Noah Lewin-Epstein, and Loraine Olson.) In this work, interested parties in education, most nongovernmental, and representing as wide a variety of interests as could be identified by examining testimony before legislative commitless on education bills, were given an opportunity for input to be used in modifying the survey design and instruments. A similar approach has been discussed by Anthony Bryk under the rubric of research design aided by stakeholder inputs. The same general orientation is evident in the current call for inputs by NCES, in which not only research investigators but also a wide range of groups with interests in education has been asked for input.

It would be wise to institutionalize such procedures for all research engaged in by NCES. If appropriately incorporated into research design, such a process can be very valuable, for neither research investigators nor government officials are in the best position to know what the emerging problems in education are. Appropriately institutionalized, such procedures become an important part of democratic processes in educational policy-making.

3. In the plans for HS&B, it was proposed to NCES by the contractor to establish an on-line HS&B data base, to make possible direct and immediate access to the data base. The data base was to be maintained either at the contractor's central

computer or that of NCES, and accessed via the educational computer network, EDJNET. Such an arrangement would have been especially valuable for those potential users whose problems, resources, and time were too small to justify obtaining public use tapes and going through the lengthy process of getting the data up, running, and able to deliver output.

NCES did not accept this proposal for dissemination and public use of HS&B data. Yet it is clear tht the time is at hand or very near for doing something like this with at least some NCES data bases. The hardware, software, and communications services are in place, so that the technologically outmoded means of disseminating NCES data (limited to printed publication or mailing of public use data tapes) can be augmented by electronic access. NCES could make its data exceedingly more useful, both for research purposes and for the wide range of other purposes that education information consumers have, by putting such a direct-access system in place - or as in the HS&B proposal, having it done by a contractor for one of its more widely-used data bases.

As promised at the beginning of Part II, the points contained here are a collection of disparate points, not connected, though I regard each as important in itself. It is, however, Part I of the paper, and the increasing importance of the outside social structure for school to which it draws attention, that I want to emphasize most strongly.

ISSUES IN NATIONAL EDUCATIONAL DATA COLLECTION

Joseph M. Cronin

Introduction

The National Center for Education Statistics plays an important role for the federal government -- collecting statistics useful for local, state and federal decision-making about education. I am pleased to join the discussion of how a redesign of NCES' data program will enhance opportunities to collect productive information and, further, promote innovative strategies for distribution, use and analysis. As my comments will illustrate, I support a simplified sampling system of educational attainment. The U.S. needs standard definitions, constant vigilance against redundancy and excessive data collection, audits and verification on local school data, and continued attention to equity results as well as excellence and achievement.

During the 1960's, I was one of several Harvard University professors of educational administration objecting to the position taken by the American Association of School Administrators to oppose national assessment and any suggestion of testing or federal action to find out what children had learned. After some discussion, AASA agreed to a compromise whereby a neutral state-oriented group, the Education Commission of the States, would conduct the national assessment of educational policy. Neither state nor local school districts would be identified in any report.

Less than twenty years later, the governors and legislatures of most states insist on knowing how well the educational system is performing with results available by school district and often by school building.

As the education officer in one Governor's cabinet and the chief state school officer in another state, I grew very dissatisfied with existing national education data and data collection formats. I remain very critical of the use by top U.S. officials -- including the President and Secretary -- of SAT and ACT results which cannot and should not be used to evaluate state achievement when the tests were intended only to predict college success for individuals. The SAT scores have been made the Dow Jones indicator of educational achievement. This is the wrong scoreboard and inappropriate information to use to compare and contrast states.

Fortunately, the Department of Education has begun to search for other indicators- the NAEP, as well as longitudinal studies on high school graduates, and other state level data. This is a positive trend.

The Growing Demand for Educational Indicators

For the last twenty-five years, the American public supported increased public expenditures for more comprehensive educational and health services -- especially for those who lacked access to quality schools, clinics and hospitals.

Neither educators nor medical administrators, however, felt obligated to compile or release to the public any systematic data about the accomplishments and failures of their institutions. Now the public wants to know, employers need to know, and government is under an obligation to collect and present these data in understandable and responsible formats.

Recently, the U.S. Department of Health and Human Services National Health Council voted to require every hospital to produce a profile of statistics which would be available to Medicare patients and the general public. Among the more compelling statistics would be the number of persons who die each year in each department of the hospital and from what ailment. Information on costs, admissions practices, length of stay, quality of staff, and a review of facilities are indeed vital statistics for hospital service consumers. Information on deaths per doctor will not be available due in part to very strong objections from the American Medical Association. Taxpayers and people seeking medical services will have considerable information due to this "truth in healing" policy and the help of professional review organizations which will monitor the hospitals.

Concurrently, an industrial location advisory service called FANTUS regularly informs companies looking for a new facility about the quality of state and local educational systems. An employer must know the literacy level of the local workforce from which new workers will be recruited. Also, employees with families transferred to the new site want to know the answer to the queries, "How good are the schools?" and "Where are the best schools?" FANTUS collects and summarizes all the available data on school size, costs, achievement, and other potentially significant indicators of quality, including the numbers of graduates going to college.

Thirty-six states have required local schools to systematically test school achievement for various reasons including grade to grade promotion and graduation from high school. New York State requires a series of tests of pupil performance evaluations well in advance of the Regents exams for college-bound students.

The Limits on SAT/ACT Data

The President of the United States and NCES ought not to try to make SAT (or ACT) scores the basis for state-by-state comparisons for these reasons:

1. The Scholastic Aptitude Test is not a test of commonly taught or needed skills. The verbal test is a cluster of reading comprehension paragraphs (commonly required skill) and a test of abilities to recognize antonyms and find analogies (which are much less commonly used in classrooms, in writing, or in life). The SAT scores, when compiled with high school grades and rank in class, do appear to predict success in college courses Freshman year. They do not purport to measure achievement in a variety of skills or subjects.

2. The SAT yields only verbal and math scores. It does not, even in the verbal segment, measure writing, speaking or listening skills. The math is somewhat more comprehensive but there are separate College Board achievement tests in algebra, geometry, calculus, and computer sciences beyond what the SAT test tries to measure. In addition, the SAT does not evaluate aptitude in science, foreign language, history, health and other important subjects or skill clusters.

3. A different percentage of high school students take the test in each state. For example, two out of three students in Connecticut (63 percent) take the SAT exams; only one in three Texas students (33 percent) take the test. These variations make comparisons, let alone rankings of the state or conclusions about quality, impossible. The tests were designed to make judgments about each student, not their local schools and certainly not state school systems.

4. State SAT scores, which are now ranked, ignore the demographic composition of each state. For example, New Hampshire, with high student scores on the SAT's, has one of the lowest number of minority students in the nation. The 1.3 percent minority are mainly associated with the air base at Portsmouth where the median education level for minority parents in the service is actually greater than that of the white adult population of New Hampshire. Also, the percentage of handicapped students of New Hampshire is only three-fifths that of Massachusetts or New York. New Hampshire has no large cities, few minorities, and fewer than average handicapped. This may explain why, despite higher than average teacher-pupil ratios and low state support for schools, New Hampshire schools and academies produce comparatively high scores. State SAT scores for New Hampshire include not only local public high schools but large national college prep schools such as St. Paul's School (Concord) and Phillips Exeter Academy, almost all of whose students take and often score very high on the SAT test.

The SAT was never designed to test statewide goals or provide state-by-state data. What is unacceptable is for the President of the United States to set state target scores on either the ACT or SAT aptitude tests since they are not achievement tests at all but specialized problem-solving measures useful for college admissions decisions.

The "Performance Outcome" section of the State Education Statistics (January, 1985) prepared by the Department of Education planning and evaluation service should be discontinued. Let the College Board (and ACT) release trend data as they do each September with the full notes about cautious interpretation. Cancel the publication of a simplistic Department of Education wall chart that does not suggest the limits of the test instrument, and which invites spurious comparison of unlike scores by the states. This chart serves poorly both college admission and national assessment movement.

What Should NCES Collect and Report?

NCES wisely has begun the search for alternative indicators, has asked dozens of researchers and organizations for advice, and issued contracts to at least one university evaluation center. This extensive consultation is prudent and the ideas produced should be useful.

I would like to suggest certain principles to guide data collection activity:

1. The data should be useful to federal, state and local policymakers or decision-makers. Much of educational research and assessment does not immediately suggest or lead to the development of a course of action. However, data on dropouts, bilingual education graduates, college-bound or job-bound students by occupation is very useful.

2. The amount of data and number indicators should be limited to that which can be stored and analyzed within three months and reported to policymakers within the year. Local and state school systems report much data to the federal government each year already that is stored but not summarized or used for other than formal report compliance purposes.

3. Data should be drawn from sample rather than total populations. The information ordinarily will be just as useful and the cost of data collection, especially to local educators, will be dramatically reduced.

Other analysts/contributors of advice to NIE will explain how the state of testing can now produce unlimited analyses of pupil performance on a thousand measures. But these important questions should be asked: Who needs it? Who will use it? For what purposes? These are deliberately hard questions, and they should be raised repeatedly about the entire program of data collection.

The state profiles and "National Report Card" should reflect a consensus among state level educators as well as educational philosophers, psychologists and psychometricians about what is worth knowing about the schools. State agency data collectors already have a strong sense of what information now is not used, or used very little or erroneously. Heed their advice. They, each and all, collect many reams of reports and tapes from local schools; they know the costs and can suggest which, if any, are beneficial.

What would be useful for policymakers to know is student achievement in grades 4, 8, and 11 or 12 in these subject skills:

- Reading
- Writing - not proofreading but a composition test
- Computation
- Computer Skills - from keyboard to programming
- Listening
- Foreign Language
- Physical Education and Fitness
- Music
- Art appreciation
- Personal health and safety
- History and government
- Citizenship

Also, it is important to know these indicators on a state by state basis:

- College (or postsecondary) intentions, and actual enrollment
- Job placement, including Military
- Dropout rates - by age and grade (with a common definition)
- Handicapped student enrollment
- Bilingual student enrollment, and length of time in a program.

The Secretary's report on excellence, A Nation At Risk, emphasized the "new basics" which were really the traditional academic subjects of reading, writing, mathematics, science, and only one "new" secondary school subject -- computer studies.

The Secretary's recommendations on the curriculum were too limited. For several decades, the Council on Basic Education has included art, music and foreign languages (all very traditional subjects associated with the cultivated, civilized, well-educated person) in their definition of basic history. NAEP included these subjects in the early assessments. The College Board and the New York Board of Regents provides achievement tests in those areas. So should any rigorous, systematic, national evaluation of education in the United States.

One caveat about computers. Technology may be a more appropriate topic. Ernest Boyer, in his volume High School, points out the limits to teaching skills or computer languages that may be obsolete in five years. More so, he suggests that the study of machines, of systems, of the history and limits of technology is of more enduring importance. If so, work might properly begin on the outline of an evaluation strategy for the 1990's in a world of fibre optics, laser technology, artificial intelligence, and genetic engineering in which computers play a major role but not the only one worthy of systematic study.

Finally, NAEP in the early years measured the educational attainment of young adults (early 1920's). Since some youth complete high school at night or during their military service, these data are part of the full picture and should be once again collected and analyzed.

Relevance, Quality and Utility of Data Collection Activities

Usually, the federal government collects data by preparing survey instruments which are screened by one or more committees and then sent out to local schools, often by a state education agency whose staff helps to administer and monitor federal education programs. This is typically the way information is collected on vocational education, handicapped education programs, Chapter One basic skill programs, food and nutrition programs, and other services where the state and local responsibilities are shared.

On several occasions, the task of data collection and analysis has been assigned to a contractor or center -- as illustrated by the use, first, of the Education Commission of the States, and now, Educational Testing Service for the National

Assessment of Educational Progress. Also, the analysis of graduating seniors and their subsequent decisions was contracted to the National Opinion Research Center. These are, for the most part, appropriate alternatives to surveying state and local school systems directly for data on educational achievement and family decisions.

The direct burden on local school systems to produce reports is much heavier than either researchers or federal program administrators realize. The U.S. has thousands of school districts with a few schools and a very small central office consisting of a superintendent and secretary, possibly a business manager. Weeks of work go into the preparation of periodic reports.

As a chief state school officer, I began to realize that local school officials tended to blame the state since the survey forms and envelopes indicated the forms, with few exceptions, must be returned to the state education agency. This makes sense because each state needs the information, and states, not the federal government, have the constitutional requirement to keep track of local school resources and activities. But, as I argued in an article entitled "The Federal Takeover of Education," the federal government puts up 5-10 percent of the funds and accounts for 50 percent of the data requests of local schools. Data on handicapped children and programs on vocational education are among the most complex, voluminous and time-consuming surveys to complete.

Since leaving the state education office, I have learned that local educators on occasion will meet a deadline by estimating the number of students served by a program. Rarely would anyone deliberately falsify statistics, even if the flow of dollars depended on a certain number. But the time and level of detail is difficult for small and medium-sized school districts and cumbersome for the large districts.

Other issues of technical quality relate to definitions. Student dropout rates are defined differently by the several states. It cannot be assumed that the reported numbers carry the same meaning from state to state. A technical task force of federal and state educators should compile all of the definitions, point out the contradictions and anomalies, and propose a consensus position for the state boards, chief state school officers and NCES to adopt.

Also, the federal government must convene a group which would standardize the definition of handicapped terms. At present, the same words carry different meanings in these agencies: Office of Civil Rights, Vocational Education, NCES, Handicapped Education, Health and Human Services, and Bilingual Education (which has a legitimate need to know how many students are both bilingual and handicapped). A few states, such as Massachusetts, have legislated the abolition of handicapped labels which tend to stigmatize, denigrate and to reduce teacher expectations.

However, certain federal bureaus and offices use outmoded or differing classification terms which make tabulation, comparison, and analysis quite difficult.

The issues summarized above can be addressed by these approaches:

1. Use of large and reputable data collection and analysis services for complex, longitudinal and analytical work on achievement and effects of educators;
2. Periodic verification of local and state data as reported on survey forms;
3. Annual reviews of data reduction options and possible redundancies;
4. A task force to review definitions of educational conditions, e.g., dropouts and handicapped;
5. An effort to obtain a uniform federal definition of handicapped services and programs.

Data Series Important to Administrators

In my tenure as a state superintendent, I needed to be informed about pupil enrollment trends, school closings, consolidations, and emerging shortages of teachers and, eventually, of facilities.

The data series on enrollments, school completions, percent in non-public school, length of school term, transportation trends, average attendance - these are bread and butter statistics needed to confirm or confront the conventional wisdom about school needs and trends.

For example, legislators suggest that if school enrollment is declining no new funds are needed. Yet, all over America, class sizes are dropping and costs of education rising. These are important trends for education advocates and legislative analysts to consider.

Data on race and linguistic background are important to administrators of bilingual programs and of intergroup or intercultural education.

At the college level, the information on earned degrees, placements and salaries are important to institutional planners and statewide coordinators of higher education programs.

Information on tuition, fees, and scholarship awards are much too skimpy given the importance of federal and state grant and loan programs. The Congress in 1980 found great gaps in knowledge about these higher education aid programs (which consume $7 billion dollars or almost half of the federal education budget and some funds from other agencies and departments). Subsequently, Congress created a National Commission on Student Financial Assistance to conduct studies and collect data which NIE/NCES might have reported. Data on grant recipients, their income levels, their race and sex, their completion rates, default rates - all of these are important to Congress and to the states. NCES can begin by reviewing the questions assigned to the National Commission and the reports filed by Commission staff and consultants, many of which should be collected at least on a biennial basis.

On the whole, this commentator believes that the NCES data series on finance and administration fills important information gaps about the resources made available to education and educators.

Other Measures of Educational Productivity

What data should be collected other than achievement data, test scores, the percentage passing courses or graduating from high school or college?

Considerations for the 1990's include:

1. The number of youth who participated in community service programs, required or voluntary, should be tabulated. The Carnegie Foundation for the Advancement of Teaching has advanced this proposal which some high schools have

adopted. This 120 hour requirement becomes a new "Carnegie Unit" and since NCES maintains the scorecard, it is important to add this tally of community service/ citizenship education requirements.

2. The amount of extra or co-curricular activities offered by U.S. high schools and the rate of student participation. For fifty years the textbooks on school leadership have agreed that students learn from student council service, from debate club, band or chorus, basketball and soccer, future teachers or farmers, and other school activities

3. The numbers of handicapped students who have been served in programs (such as those funded by 94-142) and who have:

    a. been mainstreamed, sent to less restrictive alternatives,

    b. graduated into either vocational or college preparatory programs,

    c. become gainfully employed or enrolled in college.

Now, one way to accomplish this is to engage in annual surveys. The other technique, potentially even more useful, is the periodic study of "high school and beyond," the systematic sampling of school graduates and their subsequent decisions about careers and continuing education.

The same approach pertains to bilingual education where federal and state policymakers need to know the answers to these questions:

a. How many bilingual students remain in school and graduate?

b. How many graduate from transitional bilingual education classes or programs?

c. What is the level of linguistic competence, both in English and in another language, of those who participate in bilingual programs for one school year or more?

Data Collection Strategies

A strong argument has been advanced for broadening the role of the National Assessment of Educational Progress to include state assessments as well. NAEP for several years has offered such a service to state education agencies. Several states, including Minnesota and Connecticut, have made extensive use of the data collection format and many other states have obtained test items or subsections of the NAEP instruments.

NAEP is federally funded, and one option is to add funds to the contract to enable all states to be included. Two of the major concerns will be:

1. The cost of this expanded survey, especially if NAEP requires the hiring of local test administrators as is the current practice, and

2. Timeliness of the reports, especially in more complicated assessments such as on writing skills.

The issues of cost and timeliness will be important ones to address. If assessment is too labor-intensive and costly, then inevitably some components important to civilization such as art and music or citizenship education will be eliminated.

Also, a report is useful to decision-makers if results are available within six to ten weeks — such as the College Board and ETS can provide to college admissions officials. According to some long-term participants in NCES and state education decision-making, the NAEP state survey data can take a year or more to analyze and report. A more timely analysis would presumably cost more money.

Does NAEP need to cost more money to obtain state results? Three tasks must be undertaken:

1. Legislators and chief state school officers should comment on the usefulness of the information produced by NAEP. Which findings or results are merely

interesting, and which provide data and analyses on which local and state decision-makers can act?

2. State education agency staff should be asked whether NAEP should hire and assign test administrators for each local assessment site. In carrying out an assessment (an inventory of their educational progress), Illinois found that the expensive practice of assigning an assessment administrator was unnecessary. Illinois assessment results without the local assessment administrators were not significantly different.

3. The speed of analysis and reporting must be timed to a decision-making cycle, to the planning/budgeting cycles of states (which vary) or to either a federal reauthorization or budget. Data must be current — which is one reason why the annual SAT score reports attract so much attention.

If not NAEP, there remains the alternative of asking the larger state agencies to agree on common achievement data collection or at least a common core of testing and evaluation activity. The Council of Chief State School Officers has begun this work and NCES/Department of Education should strongly support this effort. It is quite possible that assessment activity by a coalition of state educational agencies can be more cost-effective, more useful and more comprehensive than NAEP at present.

This argues against an early decision (1986) on a single ten-year format for state-by-state assessments. The Department has already made serious errors in publishing state SAT and ACT scores. The Department should consult with state-based groups, especially the legislators and chiefs and state boards of education. States do respond to incentives, to capacity-building grants, and to cost-sharing programs. The state appetite for assessment data has grown enormously in the 1980's, well ahead of any scientific consensus on how to evaluate education sensitively and thoroughly. This is an example of how federalism can work, a sharing of state and national resources.

What seems highly desirable is an agreement to use the energy and commitment of the Council of Chief State School Officers to agree on common indicators of educational progress. Such a consultation should include the tests of comprehensiveness, cost-effectiveness, and usefulness to state as well as federal policymakers.

DATA ON TEACHERS AND TEACHING:
OPENING THE BLACK BOXES OF EDUCATION

Linda Darling-Hammond
The Rand Corporation
Washington, D.C. 20037

October 1985

This paper was prepared for the Elementary and Secondary Data Redesign Project of the National Center for Education Statistics.

Over the last fifteen years, ever-declining funding for educational research and data collection has left us with many tempting clues but little solid evidence about what is actually occurring in the nation's schools. Annual reports on student test scores from the National Assessment of Educational Progress, the College Board, and individual states and school districts lead predictably to hand-wringing or back-patting on the part of educators and policymakers; but comparable, insightful trend data about the school and classroom conditions leading to these fluctuations is most noticeable by its absence from discourse about policy and practice. It simply is not available.

That we do not have a national ongoing system of educational indicators has become apparent during these last few years of commission reports on the status of American schooling. Armed with evidence of declining test scores, the various commissions have sought to make recommendations for reform based on analyses of the problems which have had to rely on old data, non-comparable data, and noniterative cross-sectional data. The lack of detailed, regularly collected trend data on important aspects of education--school finances and programs, teaching practices and methods, and teacher qualifications and assignments, among others--has forced policymakers and analysts to intuit the causes of educational "problems" (themselves poorly defined) and to infer from these intuitions what steps should be taken in response.

By treating the substance of schooling--what happens between the time that policymakers set budgets and mandates and the time that test companies take their thermometer scores--as a black box, we can never know which one(s) of any number of policy, practice, and environmental factors are producing the effects we applaud or deplore in the all-too-habitual cycle of educational "crisis," reform, disillusionment, and neglect leading to the next wave of crisis, reform, etc. The potential dangers of this approach to educational policymaking are exacerbated by the vigor with which state and local agencies have taken up the challenge to initiate reform, and by the public thirst for numbers to characterize educational progress. Where meaningful, defensible indicators of educational conditions are absent, anything that has been quantified will do, and these (sometimes conflicting) numbers are bandied about with reckless abandon, adding great heat but little light to serious deliberations about reform. The Secretary's wall chart of state comparisons is but one example of data that have been so misused; there are, of course, other less well-known but equally damaging examples.

In this paper, I would like to address two areas of data collection which I believe are most critical to understanding the conditions of education: data about teachers and about the content and methods of teaching. Let me begin by justifying these choices. Some years ago, studies of schooling were dominated by input-output methods, wherein gross measures of inputs (expenditures, class size, number of library books per student, etc.) were regressed on gross measures of outcomes (test scores, years of schooling completed, etc.) to ascertain what "works" to produce educational achievement. The answers were not clearcut, in part I would argue, because those aspects of schooling that most influence the interactions between students and teachers were ignored. In fact, the main product of this approach was a decade of debate over whether schools had any independent effects on outcomes at all. Nonetheless, these studies stimulated, for a brief time, large-scale data collection on the specified variables of interest, at least allowing examination of possible relationships and knowledge of trends.

For a variety of reasons, perhaps including (but certainly not limited to) the failure of the approach to explain differences in schooling outcomes, detailed information about school resources and expenditures ceased to be available at the national level after the late 1970s. Research efforts since then have focused on school and classroom level variables that seem under certain circumstances to produce changes in student achievement. Two "bodies" of this research have been labeled as "teaching effectiveness" and "school effectiveness" research. These sets of studies point to some generic features of school climate and of teaching behaviors that, in some instances--primarily in elementary schools serving disadvantaged students--seem to be associated with increased achievement on standardized tests of "basic" skills. While closer to the nexus between students and teachers, these studies have still not included in any systematic way the characteristics of teachers or the content of teaching as variables for examination.

I will not treat here the issues related to the validity and generalizability of these studies' results--there is a burgeoning and contentious literature on the subject--but will make two points concerning the implications of this research for data collection: (1) State and local education policymakers are seizing on these results as the basis for policy initiatives (e.g., school improvement programs, teacher evaluation systems based on "effective teaching" behaviors, remediation programs for students who have failed competency tests), yet we have no cross-cutting data sources for estimating the degree to which these "effectiveness" variables are present or absent across schools of various types, much less to examine the claims for their potency; and (2) these and other efforts to understand schools "from the inside out" have been limited by the lack of nationally-representative data about what goes on in schools and classrooms.

Now we find ourselves faced with at least two widely-rumored and much-accepted presumptions about the current state of education: that educational quality has declined and that there is, or will be, a sizable shortage of skilled teachers. Some analysts are even presumptuous enough to speculate that there may be a link between educational quality and the characteristics of teachers and teaching. And some policymakers are developing policies based on these presumptions. It is even possible that a better understanding of trends affecting the characteristics of teachers and teaching may provide some links between the deductive and inductive streams of research which have as yet failed to meet on a common ground. Yet these presumptions and possibilities cannot at present be tested with the kinds of data that are collected in either an iterative or longitudinal fashion across a representative range of students and schools.

The opportunity costs of continued failure to collect data that describe trends in the teaching force and in the substance of teaching, I would argue, are quite large, particularly at this juncture in history. Educational policymaking is increasingly tampering with the "innards" of schools, rather than merely fiddling at the periphery of school operations. Without an ongoing set of educational indicators that describes salient aspects of teaching, we will never be able to understand and reconcile the discrepant findings that result from either "black box" analyses or microscopic examinations of schooling in unique settings. We will consign educational policymaking to its traditional faddism in response to problems that are poorly understood.

THE NEED FOR DATA ABOUT TEACHERS

Spotty evidence about two recent trends have produced a waft of legislation concerning teacher education, certification, and compensation across the states. The first consists of data suggesting that the academic ability of those choosing to teach, and remaining in teaching after a few years, may be declining (Weaver, 1979, 1981; Vance and Schlechty, 1982; NCES, 1983). The second consists of data suggesting that the number of prospective teachers will soon be insufficient to meet the demand for new teachers, and that substantial shortages will result (NEA, 1983; NCES, 1980, 1982). Although this evidence is based on the best data currently available, the data are inadequate to firmly establish the existence or magnitude of these probable trends and are even less adequate to provide a diagnosis of the problem which could assist policymakers in formulating solutions.

Because there are potentially important interactions between the supply and quality of teaching candidates, a proper diagnosis of the reasons for observed or projected shortages is essential for policy formulation. The traditional responses to short supply of job applicants, in teaching and elsewhere, are either to raise salaries to increase the pool of individuals willing to offer their services to an occupation or to lower standards for entry, or some combination of both. The strategy followed depends on the degree to which timely warning of impending shortage allows for policy responses which maintain standards as well as the degree to which existing standards are viewed as useful predictors of job performance. In education right now, it is probably fair to say that there is little agreement on the types of policy responses which will allow maintenance of quality standards in the face of at least spot shortages or on the usefulness of the standards that currently exist. Thus, we see states both tightening and loosening certification and entry standards (sometimes both simultaneously), and making various adjustments in compensation at different junctures in the teaching career with little ability to predict how these changes will affect the supply of teachers or the quality of teaching.

At least two competing theories are now offered for the presumed declines in teacher supply and quality. Weaver (1978,1979), for example, has argued that decreased demand for teachers during the 1970s led to a decline in the quality of supply as schools of education attempted to maintain enrollments by lowering their standards. If this theory is correct, increased demand should of itself increase the supply of potential candidates, and tighter certification and entry standards should increase quality. In contrast, Schlechty and Vance (1981) argue that expansion of non-teaching employment opportunities for "traditional" prospective teachers has caused a decrease in the supply and quality of teaching candidates, as many have been lured away to other fields with greater financial and nonpecuniary attractions. If this theory is correct, the only way to increase teacher supply without lowering standards (or to maintain standards without further decreasing supply) will be to increase the attractions to teaching.

To test these (or other) theories about how the size and composition of the prospective teaching pool evolves, we need data about how occupational choice decisions are made by college students and how these decisions translate into actual employment decisions after graduation. However, to estimate the effects of various policy alternatives on overall supply and demand we also need to know about the size and character of the reserve teaching pool, the conditions under which its members will offer their services to education, the determinants of teacher attrition, and the degree to which certain teacher attributes or skills are interchangeable when shortages exist in some teaching fields while surpluses exist in others. Models for projecting teacher supply and demand must incorporate not only these kinds of data, but also take into account policy-generated changes in demand and policy-generated definitions of shortages which rest on conceptions of which characteristics of teachers are essential and which are dispensable for filling certain kinds of teaching positions.

The definition of shortage as something more than a count of unfilled vacancies is most essential if we are to move beyond a warm-body theory of teacher quality. Although measures which take into account teacher certification or college coursework preparation move us further toward some notion of supply which incorporates a quality dimension, the knotty issues of what knowledge and skills are important to teaching will ultimately require data on teaching practices matched to data on teacher preparation if we are to begin developing an understanding of how policies relating to teacher supply influence the actual content and outcomes of teaching. And, as I have argued above, until we begin to understand what actually occurs in classrooms, we will never solve the riddles posed by incomplete models of school effects which first stimulate and later dampen reform initiatives. Below I propose some of the most essential data collection efforts I believe are necessary to begin sketching out trends concerning teachers and teaching, which might someday lead to answers to perennial policy dilemmas.

Teacher Supply

Projecting teacher supply requires knowledge of at least three sources of potential supply:

1. The number of current teachers expected to remain in the teaching force at a given point in time (the required estimate of attrition is, of course, also a component of projected demand);
2. The number of college graduates expected to choose teaching as their initial occupation. (For long-range projections, we should also consider the proportion of these who are expected to remain in teaching by year X or Y);
3. The number of individuals qualified to teach who are currently not teaching but might return to teaching, i.e., the potential reserve pool.

None of these estimates is simple to derive, and currently available data sources are inadequate for each. To complicate matters further, local or regional supply estimates must take account of in- and out-migration from the labor market area. Field-specific estimates must take account of substitution possibilities among teachers in different teaching areas. Putting aside these additional data requirements for the moment, let us examine how well we can handle the basic task of projecting supply with current information.

The first component, the stock of current teachers minus attrition, ought to be the simplest to estimate and project. However, though the number of current teachers is known, current and prospective attrition rates are not. The most recent estimate of teacher attrition was obtained by NCES in 1969. This estimate of 6 percent has been used ever since in NCES projections of teacher supply and demand.

There are a number of reasons to believe that teacher attrition rates are not static. First, the age composition of the teaching force changes over time; hence, the proportion of the force nearing retirement also changes. In addition, recent data from a number of states and school districts suggest that attrition rates are especially high (50 percent or higher) for inexperienced teachers during their first few years (cf Mark and Anderson, 1985; Grissmer and Darling-Hammond, 1984; Vance and Schlechty, 1982). Thus, the experience composition of the teaching force--also related to the age-distribution--may be an important (and changing) variable. Third, labor market forces in teaching and in the general economy undoubtedly influence turnover. When teaching positions are scarce, temporary exits may be fewer due to expected difficulty in re-entering; when other opportunities are plentiful, career changes are more likely. (These opportunities may also, of course, be more plentiful in some regions of the country and for teachers in certain fields--especially math, science, and computer science--than others.)

Finally, policy variables may influence attrition rates. Incentives for early retirement, for example, became widespread in school districts during the 1970s, when declining enrollments required reductions-in-force. These incentives may now work, ironically, to produce shortages. Current policy initiatives, such as internships for beginning teachers and merit pay or career ladders for veterans, are intended to influence attrition rates. Perhaps they will.

Suffice it to say that we should expect attrition to change with the shape of the teaching force, with the health of other sectors of the economy, and perhaps even with changes in policy affecting teachers. We ought, then, to be prepared to regularly estimate attrition rates for various classes of teachers as a basis for preparing and modifying projections of extant supply.

As a first step, we need to know the age composition of the current teaching force and attrition rates by age category, so that we can project the number of retirements and early leavers over time. It is important to note that, as new teachers comprise a greater share of the teaching force in the coming years, it will be more important to understand when and why many of them leave, and whether they plan to return. Currently the NCES Common Core Data Set tabulates state counts of full-time equivalent teachers by level, but not by age, experience, or teaching field. The periodic NCES surveys of teacher demand and shortage collect data from a sample of school districts on total teaching positions by field and on vacancies and new hires, but do not allow accurate estimates of attrition. New sources of data must be developed if these information needs are to be met.

While surveys of school districts or schools could be developed to provide estimates of localized turnover rates and probably fairly accurate estimates of retirement, they would need to be supplemented by ongoing surveys of teachers if mobility among districts is to be accurately separated from temporary and "permanent" leaves from the profession--and if reasons for leaving are to be understood.

In addition, both district-level and teacher-level surveys should collect data that will allow analysis of policy factors that might influence both entry and retention in teaching. At a minimum, salary data must be collected from districts (through the CCD and/or other district surveys) that describe the salary range and distribution of teachers across that range in addition to average salaries, which are uninformative for most analytic purposes. Information on retirement plans and other important compensation variables would also be helpful for analysis of attrition rates. Cluster sampling of teachers within districts to establish career paths and plans could further illuminate links to policy variables, especially if these surveys could include information on teachers' views of the policies as they influence career decisions.

The second component of supply, entrants who are recent college graduates, is also problematic. Traditionally, estimates of new teacher supply have been made on the basis of the number of students graduating with degrees in education. Although data are now routinely collected which provide such estimates at the bachelor's and master's degree levels, and these are incorporated into NCES projections, there are problems with using this measure alone to estimate and project new teacher supply.

First, we need to know what proportion of master's degrees granted in education are awarded to persons preparing to teach as opposed to those already teaching. Previous inclusion of master's degree students in estimates of new supply have caused overestimates; exclusion will lead to underestimates of unknown proportions. While in the past, the vast majority of such master's degrees were awarded to existing teachers, this may change as states open up alternative routes to certification for liberal arts graduates and as some teacher education programs move to a 5-year model.

Second, we need to know what proportion of students receiving bachelor's degrees in education actually plan to teach, and how many do so. Prior NCES surveys of recent college graduates have obtained such data; however, they do not provide information about the degree to which failure to enter teaching is due to inability to find jobs or to changes in occupational plans. Clearly such information is needed if we are to understand the real sources of supply.

The proportions of teacher education graduates who do not initially enter teaching vary substantially over time and across teaching fields. In 1976-77, NCES estimated that the portion of newly qualified graduates seeking teaching positions was 77 percent, with only 60 percent ultimately accepting teaching positions. In 1981, the estimate of those seeking full-time positions was 85 percent, with 64 percent ultimately accepting such positions. Differences among teaching fields are also substantial. In 1981, for example, only 30 percent of prospective health teachers accepted full-time teaching jobs as compared to 75 percent of prospective special education teachers (NCES, 1983).

Previous studies have tended to overestimate new teacher supply because of lack of data about occupational decisionmaking both during and immediately after college. One consistent source of overestimates in projections has been the assumption that the teacher production rate is a constant share of the college student population. In fact between about 1970 and 1980, the proportion of college students majoring in education declined by nearly half--from 21 percent to 11.6 percent--and the proportion of college-bound students now planning to major in education is only about 5 percent. Thus, projections of teacher supply must incorporate *trends* in the occupational choicemaking behavior of students. These trend data are useful only if they are understood. To what extent are substantial changes in teacher production rates a response to labor market factors (i.e., a perceived surplus of teachers)? To what extent are they evidence of disaffection with the salaries or other conditions of teaching?

The relationship between educational and occupational plans and actions must be examined at several points in the teacher production pipeline if we are to understand the factors influencing the supply of new teachers: at college entry, choice of major, college exit, and initial occupational choice. As mentioned earlier, the decision to leave or remain in teaching during the first few years is also extremely important. The NLS supplement to be conducted next spring that will survey all of those in the 1972 sample who ever majored in education or taught (plus a sample of those who might otherwise be classified as prospective teachers) will provide a valuable data source for examining these questions for that cohort of students. NCES should consider adding a small number of carefully-designed questions to the HSB followups of 1980 sophomores and seniors to track their occupational decisionmaking with respect to teaching as well.

The third source of teacher supply--the potential "reserve pool" of teachers--is most difficult to estimate, but may be increasingly important if the number of new college graduates entering teaching continues to shrink. These are individuals who are qualified to teach but who are not currently teaching, either because they, perhaps temporarily, left teaching to raise families or pursue more education, or because they have entered other occupations. Estimating the size of the real reserve pool depends on knowing what proportion of these individuals would consider re-entering teaching under various conditions as well as knowing the annual rates of entry into and exit from the pool. Previous studies have tended to overestimate the annual supplies from the reserve pool since an overprojection of the number of new teachers produced in any year results in an overprojection of the size of the reserve pool in subsequent years and because "attrition" from the reserve pool (into other occupations) is not taken into account. Even with better estimates of annual supply, we will need data that allows us to identify the factors that influence the behavior of exteachers and their relative propensities to seek teaching positions.

There are several possibilities for collecting such data; though each is imperfect, they would provide us more information than we now have. The special NLS survey of current and former teachers will provide some information on the current occupational status of former recruits to teaching and, perhaps, on their plans and attitudes toward re-entering teaching. Since this is a single cohort, though, it will provide limited information on responses to different labor market conditions and on attitudes of later cohorts. Nonetheless, the survey should provide strong indications of the numbers and characteristics of the "potential" reserve pool for that cohort who are in fact firmly committed to other occupations and lost to teaching. Similar followups with the HSB sample would strengthen the analysis of reserve pool behavior. In addition, ongoing surveys of teachers might be designed to solicit personal and job histories from those who have just entered or re-entered teaching and job plans from recent attritees (if they are tracked) or those who have definite plans to leave during the school year. These efforts would help in developing estimates of both entry and exit rates from the reserve pool for different types of teachers and an understanding of the factors influencing these decisions.

Teacher Demand

To estimate and project teacher demand we need to know the number of teaching positions required in a given year (by field and level), and the number that will be filled by teachers currently in the teaching force. The first component, the size of the teaching force, depends at the most general level on student enrollments and anticipated pupil/teacher ratios, both of which are fairly easy to estimate with existing demographic and school data. Projection errors, though, can stem from at least two sources:

1. Inability to anticipate student enrollment growth due to such factors as immigration (which may cause dramatic increases in enrollment in some regions or localities) or changes in student service patterns, caused by policies that extend public education downward to kindergarten and preschool levels as is occurring now in some states.

2. Inability to anticipate policy-generated changes in staffing patterns, such as those which accompany new forms of service delivery (e.g., the now fairly widespread use of specialists in elementary schools or the possible changes in staff responsibilities that may accompany career ladder plans); or new programs and course requirements for students (e.g., the addition of special education, bilingual education, and compensatory education programs during the 1970s, or the changes in student graduation requirements now being enacted in many states).

These kinds of changes certainly affect the demand for particular types of teachers; depending on how schools manage these changes, i.e., the degree to which they substitute or supplement teachers and services), they affect the total demand for teachers as well.

There are at least two possible means for improving sources of data about these elements of teacher demand. First, it may be possible to include questions about recent policy changes presumed to affect teacher demand in the CCD surveys of state education agencies and in the Demand and Shortage surveys of local education agencies. While officials may not be able to estimate the effects of recent policy changes on teacher demand, the availability of other state- and local-level data about student participation in the relevant courses, programs, or services might allow analysts to do so.

At a minimum, though, collecting such information would allow some gross adjustment of demand projections if used in conjunction with a second source of data: regular reports of teacher vacancies in particular fields *which separate new demand from demand produced by attrition*. The periodic surveys of teacher demand and shortage are not currently designed to provide this information, although the addition of a few questions would allow them to do so. Specifically, the surveys need to ascertain the number of teaching positions (by level and field) authorized for the current year *as compared to* the number authorized and filled in the previous year (now not asked), along with the number of continuing teachers, new hirees, and unfilled vacancies. Adding questions about student enrollment trends and pupil/teacher ratios would also allow much more accurate understanding of the sources of demand.

The second component of demand, the number of continuing teachers, was treated in the previous section's discussion of teacher attrition. One additional observation is important here. The degree to which teachers with different skills and preparation are substitutable when new areas of demand emerge has important implications for assessing both overall and field-specific demand and shortage. Much of the disagreement over current shortage projections stems from lack of clarity on this point.

Teacher Shortage

As an example of how definitions of teacher shortage vary, consider two extreme views of teacher fungibility. On the one hand, if teachers of different backgrounds are always perfect substitutes for one another (e.g., an elementary school teacher can as easily teach high school mathematics and a junior high industrial arts teacher can as easily teach kindergarten), then estimates of total teacher supply and total teacher demand are all that count. If supply equals or exceeds demand, and is projected to continue to do so, there is little cause for concern about shortages. (Obviously, local and regional imbalances in total supply and demand are important and will vary. Some excess in supply is always necessary to force mobility to places with higher demand and to allow quality distinctions to be made in hiring.)

In fact this is the basic approach of most general projections of teacher supply and demand, which further assume that anyone who is teaching (or has taught) can be counted as a teacher, i.e., part of the supply pool, regardless of qualifications. This assumption stems from the lack of agreement about what constitutes qualification to teach, leading to non-discriminating measures for counting teachers that reflect states' and school districts' willingness to hire individuals without standard credentials as teachers when the need arises or to reassign current teachers outside their areas of preparation. These then become part of the teacher pool, and it becomes almost impossible to discern a shortage, since vacancies are nearly always filled somehow with someone. The analogy would be to calculate in the supply of physicians anyone willing to offer his or her services as a doctor, regardless of training or licensure, in an environment where significant bars to this practice did not exist. Thus, the recent surveys of Teacher Demand and Shortage report few "shortages" as measured by unfilled vacancies, while projections based on qualifications to teach have anticipated shortages and surveys of teachers suggest that a nontrivial proportion teach outside their fields of preparation or certification.

At the other extreme, if we assume that teachers are totally non-fungible and that particular preparation--however defined--is essential for successful performance of the job, i.e., that a teacher can only teach those subjects or levels for which they have particular preparation, then supply, demand, and shortage would be assessed much differently. Separate computations of supply and demand by field would be essential, with acute shortages obvious in some and surpluses obvious in others, and individuals hired or assigned to teach in areas for which they lack the particular qualifications used as a discriminating measure would be eliminated from counts of "legitimate" supply, thus producing measures of latent if not blatant shortages.

Obviously, taken to the extreme these assumptions can become equally nonsensical, and NCES cannot become the final arbiter of teacher quality measures. However, some indication of the degree to which the demand for specific types of teachers is matched with a supply of appropriately-trained teachers is essential for policymakers concerned with teacher supply and quality. Given that there is disagreement about measures of teacher quality, several different measures could be used in data collection efforts that describe the stock of teachers and their teaching assignments. These might include certification in the field(s) taught, college coursework preparation in those fields, and pedagogical preparation. Such indicators would at least allow policymakers to track supply, demand, and shortage according to various definitions of "legitimate" supply. They would also allow some means for reconciling currently disparate estimates and, ultimately perhaps, for examining how school districts' hiring and assignment practices influence other teaching variables of interest.

Although data are not available to demonstrate conclusively how qualifications-related measures of supply and demand would affect estimates of shortages, some sense of the possible magnitude of differences in estimates derived from alternative assumptions can be gained from recent surveys. The 1983-84 NCES estimates of teacher shortage, based on a measure of unfilled vacancies reported by a sample of school districts, indicate overall shortages in the neighborhood of only 1.6 per thousand current teachers, with field-specific shortages ranging from .4 per thousand for reading to 8.8 per thousand for bilingual education (NCES, 1985). This range probably reflects supply as it interacts with the outer bounds of teacher "fungibility" in different fields. That is, a number of individuals might well be viewed as capable of teaching reading, but the potential supply of bilingual education teachers is limited to individuals who are, in fact, bilingual themselves, aside from the application of any credentialing standards.

Applying a different standard leads to quite different estimates of shortage. For example, the same 1983-84 survey provided estimates of the proportion of total and newly hired teachers not certified in their principal field of assignment; these amounted to 3.4 percent of all teachers and 12.4 percent of all newly hired teachers. If we assumed that no certified applicants could be found to fill the vacancies filled by newly hired uncertified teachers and added these 26,300 positions to the count of unfilled vacancies, the estimate of shortages would increase dramatically from 1.6 to over 20 per thousand current teachers. If we further assumed that the positions filled by other teachers assigned outside their fields of certification could not have been filled by certified applicants (for various reasons this becomes a more dubious assumption), the estimates would skyrocket further.

Applying still more rigorous standards yields predictably larger estimates. Since certification is not a perfect measure of preparation,[1] we might want to know what proportion of teachers are teaching classes outside their fields of preparation that might otherwise be counted as evidence of demand or, if unfilled, of shortage. (Again, this requires inferences about hiring and staffing and disallowances of substitutions or economies that are not entirely supportable.) A 1980-81 NEA survey of teachers indicates that 16 percent of all teachers teach some classes outside their field of preparation; and 9 percent spend most of their time teaching "out of field" (NCEA, 1981); the HSB survey indicates that, among high school teachers, 11 percent teach primarily outside their area of state certification and 17 percent have less than a college minor in the field they most frequently teach (Carroll, 1985).

To be sure, we do not know the degree to which such "out of field" assignments are actually inappropriate according to various standards, or the degree to which they impair teaching quality; nor do we know the extent to which the discontinuation of some of these types of hiring and assignment practices would actually influence teacher demand or shortage. Some of these practices are undoubtedly the result of capitalizing on teachers' individual uncredentialed abilities and interests; some result from district attempts to continue to employ senior teachers when demand in their particular teaching fields declines; some are probably supported by inservice training that upgrades the stock and flexibility of human capital. On the other hand, misassignment as a response to teacher shortages may also result in poor teaching in some unknown proportion of instances, and may contribute to teacher stress and attrition.

In order to understand the interactions between teacher supply and qualifications and the effects of qualifications on teaching practice, we need data from ongoing surveys of teachers to supplement the data on teacher hiring and vacancies collected from school district personnel offices. The latter are useful for providing gross measures of shortage and qualifications, but cannot provide details about teachers' characteristics and attributes or about teaching practice. Through the preceding discussion, I have described certain kinds of information needed to understand important aspects of supply and demand that can only be collected from teachers, and must be obtained on a regular basis. Ongoing surveys of teachers such as those launched in the Public and Private School Teacher Surveys of 1984-85, should be designed to collect data on teachers formal qualifications, additional inservice training; teaching assignments and job histories and plans. Furthermore, as described below, these surveys should collect data on teachers' attitudes and teaching practices so that policy-related questions concerning the links between school and teacher variables can be examined.

---

[1] The HSB survey of high school teachers, for example, indicates that of the small number (1.7 percent) of high school teachers who have had *no* college courses in the field they most frequently teach, 74 percent are nonetheless certified in that field (Carroll, 1985).

THE NEED FOR DATA ON TEACHING

Currently, very little data are available on teachers' working conditions, their views of school conditions and administrative practices, or on the activities they engage in as part of their instructional efforts. The only ongoing source of information providing trend data on teachers' working conditions, career plans, and views of teaching is collected every five years (since 1956) by the National Education Association, in the Status of the American Public School Teacher surveys. For the first time, the 1985 NCES Public and Private School Teacher Surveys asked questions about teacher qualifications, teaching load and salaries, but there are no comparable questions about career plans and satisfaction with which to test links between working conditions and career decisions. Future surveys of teachers conducted by NCES ought to seek data on all of the these variables on a regular basis so that possible relationships between teacher qualifications, compensation and work load, views of working conditions, and career decisions can be examined.

Still less data are available about what teachers actually do in classrooms, much less how teaching practices relate to teacher attributes or school policies. Though studies of school and teacher effectiveness have suggested a number of policy and practice variables that are important to schooling outcomes, no nationally-representative source of data on these factors is available to examine trends or relationships. The addition of teacher and administrator questionnaires to the HSB student surveys is a commendable start for understanding the experiences of high school students. NAEP's efforts in this direction will also enrich our understanding of schooling processes for the NAEP samples of students, teachers, and administrators. However, the usefulness of these data for various policy purposes is limited by the nature of the sampling frames appropriate to their major purposes.

A tantalizing peek at changes in teaching practices was provided by the major longitudinal studies data on high school seniors reports of teaching methods used in their classrooms. Between 1972 and 1980, students were exposed to writing, student-centered discussions, and project or laboratory work in their public school 12th grade classrooms. More were exposed to individualized instruction and computer-assisted instruction. Unfortunately, the data did not allow further examination of teaching practices or their relationship to teacher attributes or school policies. And, of course, no such data are available for examining teaching practices at other grade levels.

Among the major questions about teaching for which we have no trend data available are the following:

- How much instructional time is available to students and teachers (e.g., length of the school day and school year; duration of classroom periods)?

- How do teachers use instructional time (e.g., amount of time devoted to different subject areas and to different teaching activities such as lecturing, discussion, seatwork, project or lab work, recitation, testing, reading, writing)?

- How much instructional time is lost to "nonteaching" activities (e.g., paperwork, pep rallies, class changes)?

- How much emphasis do teachers place on different instructional tools or methods in classroom and homework assignments (use of textbooks, workbooks, computers, teacher-developed materials, library books, research projects or other problem-solving activities, writing themes, etc.)?

- How do teachers make decisions about what and how to teach district or school policy, textbook coverage, test coverage, personal views of student needs, etc.)?

Data on teaching variables such as the above, if gathered from teachers sampled in clusters by school and district and, if combined with administrative data on policies and school/student characteristics, would allow analysis of how policy and environmental factors influence teaching practice; how teacher characteristics--including qualifications and experience in the teaching area--influence practice; and how practices change over time. It is beyond the mission of NCES to establish how any or all of these variables affect student learning--other kinds of research and data collection activities sponsored by NIE and elsewhere must attend to these questions--but as knowledge grows about important indicators of effective schooling and teaching, NCES should strive to incorporate sources of relevant data in a stable, ongoing system of teacher and administrator surveys that reveal something of what actually occurs in those black boxes called schools.

This effort will require greater amounts of resources than have been available to NCES in recent years. However, the resources required are negligible compared to the millions of dollars expended on the thousands of commission meetings and papers and other activities attendant to reform formulation which sweep the country each time a new crisis in education is declared. Though NCES cannot relieve us of crises and commissions, it can provide a basis for deliberations about problem sources and solutions which might advance the debate, and perhaps eventually break the cycle of educational crisis, reform, disillusionment, and neglect which is repeated with such distressing regularity.

REFERENCES

Carroll, C. Dennis, *High school and beyond tabulation: Background characteristics of high school teachers*. Washington, D.C.: National Center for Education Statistics, January 15, 1985.

Grissmer, David, W., Linda Darling-Hammond, and Corazon Francisco. *A prototype personnel planning system for the D.C. public schools*. WD-2547-DCPS. Washington, D.C.: The Rand Corporation, February 1985.

Mark, Jonathan H., and Barry D. Anderson. Teacher survival rates in St. Louis, 1969-1982. *American Educational Research Journal*, 1985, *22*(3), 413-421.

National Center for Education Statistics (NCES). *Projections of education statistics to 1988-89*. Washington, D.C.: U.S. Government Printing Office, 1980.

National Center for Education Statistics (NCES). *Projections of education statistics to 1990-91*. Washington, D.C.: U.S. Government Printing Office, 1982.

National Center for Education Statistics (NCES). *The condition of education, 1983 edition*. Washington, D.C.: U.S. Department of Education, 1983.

National Center for Education Statistics (NCES). *The condition of education, 1985 edition*. Washington, D.C.: U.S. Department of Education, 1985.

National Education Association (NEA). *Teacher supply and demand in public schools, 1981-82*. Washington, D.C.: NEA, 1983.

Schlechty, P.C., and Vance, V.S. Do academically able teachers leave education? The North Carolina case. *Phi Delta Kappan*, October 1981, *63*, 106-111.

Vance, V.S., and Schlechty, P.C. The distribution of academic ability in the teaching force: Policy implications. *Phi Delta Kappan*, 1982, *64*, 22-27.

Weaver, W.T. Demography, quality, and decline: The challenge for schools of education for the 1980s. In *Policy for the education of educators: Issues and implications*. Washington, D.C.: American Association of Colleges for Teacher Education, 1981.

Weaver, W.T. In search of quality: The need for talent in teaching. *Phi Delta Kappan*, 1979, *61*(1), 29-31, 46.

Weaver, W.T. Educators in supply and demand: Effects on quality. *School Review*, August 1978, 552-93.

IMPROVING THE QUALITY AND UTILITY OF NCES DATA

Jane L. David
Consultant
3144 David Avenue
Palo Alto, CA 94303

21 June, 1985

Introduction

The National Center for Education Statistics (NCES) seeks to redesign its data collection efforts to improve the quality and utility of the data for decisionmakers and the general public. For this to occur, three goals must be met. First, the choice of what data to collect must be driven by the questions of interest to decisionmakers and the public. Second, procedures must be in place that insure the data are valid and reliable. Third, the data must be reported in ways that facilitate use by the intended audiences.

For NCES, the second goal is the most critical. Careful choices about what data to collect and clear reporting cannot compensate for inaccurate data. Below I first comment on data accuracy and then discuss reporting, citing examples from Indicators of Education Status and Trends (January 1985). I then list specific comments on Indicators of Education Status and Trends with page references.

Data Accuracy

The biggest challenge facing NCES is that of insuring the validity and reliability of the data they report. If the data continue to be as inaccurate in the future as they have in the past, all other issues are moot. The more levels of aggregation the data pass through, the more sources for error. Because NCES must rely on second, third, and fourth hand data, it is essential to put into place a set of procedures designed to check the validity and reliability of the data.

Given the need to rely on data from other sources (particularly state administrative data which are notoriously inaccurate), NCES must, at the least, develop a system that permits crosschecking the data with other sources for the same information. Judging from the description of current data sources, it appears that there are multiple sources of data for certain types of information (e.g., data on staffing and teacher characteristics reported by the states are also collected in NCES's Public School Survey). To the extent that multiple data sources already exist, NCES should make comparisons across data sources and report on both the extent to which discrepancies are found and plausible explanations for the discrepancies. NCES should exploit all opportunities to corroborate data sources over which they have little direct control, such as state administrative data and Census data.

Data for which multiple sources do not currently exist should be collected through alternative means designed explicitly as a crosscheck. The Fast Response Survey System could readily be adapted to this end. Of course, this presumes that care is taken to ask precise questions. Unfortunately, my only experience with NCES Fast Response Survey data suggests that these data are often inaccurate as well. On the one occasion in which I used these data (1979), my own telephone surveys corroborated by field work produced figures quite different from those of NCES (e.g., NCES reported 66 schools in California participating in the Schoolwide Projects Provision of what was then ESEA Title I; I located 107 such schools).

Perhaps NCES could also establish samples of schools within states (if such a sample does not exist) as an extension of the Fast Response Survey System. NCES could also build items into contracted longitudinal studies for purposes of corroborating other data sources. In addition, NCES should be aware of other national data collection efforts (particularly federally funded studies in education, labor, and health and annual surveys such as Gallup, Louis, etc.) and develop agreements for sharing data with the funding or data collection agency.

Reporting

Given the nature of the data to be collected, and the reliance on indirect sources and multiple levels of aggregation, there will always be issues of validity and reliability. However, NCES can take steps (a) to maximize validity and reliability and (b) to inform readers of the weaknesses that remain.

In reporting the data, it is absolutely essential to have indicators of the validity and reliability of the data. As a user of data, I am always suspicious of any type of survey data or compilation across levels of government. However, when I know how the question was asked, I can draw my own conclusions about the bias of the responses. When I know the sample size and a standard deviation or a confidence interval, I can draw my own conclusions about its credibility and utility. Through television and other media, even lay audiences are accustomed to confidence intervals and other indicators of measurement error.

Based on a close reading of the Indicators of Education Status and Trends, I urge NCES to consider the following recommendations for reporting the data in addition to reporting how questions were asked and estimates of reliability.

1) Following the previous recommendations for corroborating data, reported data should include a brief description of the similarities and discrepancies from different sources and an analysis of what accounts for the differences. To the extent that the discrepancies influence interpretation of the data, the text should alert the reader to the limits on interpretation.

2) When multiple sources are cited, as is currently the case for many of the tables and charts in Indicators, indicate how the multiple sources were used. Multiple sources always suggests to me that data were merged without regard to comparability. When multiple sources are cited, note which data came from which source and, if two sources were merged for one estimate, describe how and why this was done.

3) Precise descriptions of what the numbers represent and interpretations of how their known weaknesses influence the results is crucial. Some inaccuracy is inevitable in such a mammoth system of data collection. No one expects otherwise. The utility of the data rests on attaining an acceptable level of quality and alerting the reader to its remaining weaknesses. Unfortunately, the text in Indicators accomplishes neither of these goals. Sloppy table titles and category labels are inexcusable as is vague and ungrammatical prose. (See examples below.) Readers must understand the limits of the data.

4) Knowledgable interpretation of data is useful to readers; uninformed or sloppy interpretation is dangerous. NCES should provide interpretation but not without a system that corroborates the interpretation(s). One approach is to create panels of outside reviews in different areas (e.g., one for student performance, another for human resources) who would comment on a draft of the presentation. Agreement is not essential; presenting conflicting interpretations is also extremely useful to readers.

5) Collect and report the data in ways that minimize inappropriate comparisons. For example, comparing states by comparing trends over time within states decreases the problems due to different measures and definitions. NCES does this well in several places. In addition, the text around tables and charts should draw the readers' attention to the appropriate comparisons.

6) NCES should cite data sources more accurately and fully. When "NCES estimates" are cited as the data source, my eyebrows go up. All sources should include dates.

7) Use the glossary to help readers. Dictionary definitions of enrollment, attendance and biology, for example, are not helpful. The Glossary should reflect the way questions were asked and indicate differences in definitions across states or data sources.

Comments on "Indicators"

The concept of a report on indicators of education status and trends is excellent. This kind of annual report has the potential to provide an invaluable picture of our educational system over time. At the least, it provides a backdrop against which to interpret other data and educational issues at all levels--national, state and local. At the most, the data can inform the public and policymakers about expected trends in such critical areas as teachers supply and student enrollment.

Recognizing that the January 1985 Indicators of Education Status and Trends is a first attempt which seeks reactions from potential users, I offer the following specific reactions by report section and page.

Outcomes

p. 3 I find the NAEP data by assessment area within subject particularly illuminating. I realize presenting subscales greatly increases the quantity of the data. This is an instance in which some clear text around the tables would be useful. The text could report conclusions from an inspection of more detailed tables and reference other publications. (A minor point: describing shifts in performance "over the past decade or so" is misleading, especially for science which includes no data from the past eight years.)

p. 6 I don't know what to conclude from these data. How were these topics measured? Did the items correspond to a particular curriculum? Did they emphasize computation or problem solving? Are the numbers medians of 18 means? Do the results look different if the comparison is between the United States and the five countries with the same proportion enrolled in math?

p. 8 Because this is an area in which many states are changing their requirements, it would be useful to see shifts over time and a breakdown by state. Are these only comprehensive high schools?

p. 10 This is an instance in which how the question was asked and to whom is critical. Can GEDs be reported separately? (It isn't clear whether they are included on p. 10 or not).

Resources

p. 24 Basing pupil/teacher ratios on all instructional staff results in a gross underestimate of class size. Readers will draw from this table an image that severely distorts reality. These data must be presented for teachers with regular classroom assignments or not at all. There is no indication of what the sample is for the class size data; are they based on regular teachers only?

Why is the research literature mentioned here and not elsewhere? If research is mentioned, there should be complete citations. I don't think it belongs here; an adequate summary is difficult to do in one sentence.

"In 1971 and subsequent years, the data by level are estimated." The footnote should say how and why the estimates were made. Citing "unpublished data and estimates" only invites suspicion.

p. 26-27 The headings and labels for the table on p. 26 are a good illustration of the need to use precise language. The title is extremely misleading; this is not a table about the distribution of academic ability in the teaching force. Throughout the present report, typically the charts on the right are far easier to understand than the tables on the left. This is partly because graphs are often easier to grasp, but more due to the clarity of the titles and headings. The title and labels of the graph are much more precise and hence clearer than those on the left. (E.g., Percent Scoring in Highest Fifth is much clearer than Highest Rank with a footnote saying the sample was broken into five ranks).

The content of these two pages also raises questions. First, it is unconscionable to refer to these data as measures of the "Quality of the Teaching Force." No single test score can capture teacher quality. Moreover, given the small percent of teachers who take the SAT (and no indication of this in the table), I conclude that the numbers are extremely misleading. I have no problem with the idea of looking at teacher performance on academic measures, but I need to be convinced that these data speak to that issue. Perhaps this is an instance in which the National Longitudinal Study is not the best source of data. (The samples seem quite small.)

p. 28 For data about teacher supply and demand to be useful, they need to be reported separately for elementary and secondary school and by subject area. It would also be useful to see these figures by region and the same breakdowns for teachers "teaching out of subject"--that is, teaching in fields for which they are not certified. The table doesn't state that the entries are in thousands.

Context

p. 37 Do data on teachers' perceptions of problems exist over time? The wording of the second bullet suggests that the Metropolitan Life/Harris Survey was conducted in other years as well. It would be useful to see trends in these data. Were the questions posed to the public and to teachers in the same way? If not, since the tables invite comparisons, it would be useful to know how they differ.

p. 38 Certainly school environment indicators are of interest. However, the key factors associated with effective schools are not amenable to the kind of measurement that could be aggregated across schools and districts and states. Most of us don't even know how to measure them on site. The kinds of perceptions to be reported in the 1984 follow-up for High School and Beyond may prove useful. I suspect that the greatest utility will lie in the items that have been asked over time. As with trends within state, shifts over time are easier to interpret than absolute levels of factors like "environment conducive to student achievement." This suggests a critical need for ongoing longitudinal studies of this type.

p. 40 How many students actually receive these various types of services?

p. 42 Where are the data cited in the text on pages 42 and 43? To estimate school enrollment trends, are census data available on numbers of babies/children ages 0 - 3?

p. 44 Referring to state required Carnegie units as "State Governance" seems odd.

Additional data that I would find useful include:

- median age of teachers by state
- data on preschool attendance
- measures of student mobility/turnover
- information on number of hours worked (and types of jobs) for high school students by state and minority status

Looking to the Future

The proliferation of microcomputers in district offices and schools, combined with growing sophistication about and access to telecommunications, has far reaching implications for future data collection. Now is not the time to implement such a system; neither access to the technology nor user sophistication is sufficiently widespread. But now is the time to begin to design a computerized data collection system utilizing the telecommunications capability that most districts and schools will have within a decade or less. Such a system will require considerable planning and testing; waiting until the technology is completely in place will put NCES a decade behind. If General Motors can design a system that automatically translates an individual customer's order into instructions for what parts to manufacture and into a custom made car, NCES ought to be able to gather basic descriptive information about our schools through similar applications of technology.

The second application of technology that NCES should now be investigating is the use of microcomputers for different kinds of assessment instruments. The limits of paper and pencil tests are well known. Designing new measures that go beyond simple multiple choice questions should be underway.

Eugene E. Eubanks
School of Education
University of Missouri-
Kansas City
5100 Rockhill Road
Kansas City, MO 64110

Eugene E. Eubanks

As we proceed into the 1980's and into the 1990's, our public schools in the big cities or urban areas will become even greater proportionally minority and poor. The needs, aspirations and ability of this poor and minority population in our big cities will require the development of new sources for data collection because the traditional measures used to collect data on this population have proved to be most inaccurate. Inaccurate data on this population has far reaching consequences for the future of America when they lead to policy decisions which address problems that no longer exist, problems that are not adequately defined and, in too many cases, problems that never existed. As well intended as these policy decisions might be, our country, with resources now scarce, cannot direct and use these resources by chance and in far too many instances use them to aggravate problems.

What is the data which will prove to be of the most value to persons working in the trenches at the public school level? Well, in my personal experience and knowledge gleaned from researchers, the following items are suggested with the idea that such data might lead to some reform. The fact that such data are collected could lead people to pay attention to them.

1. The NCES should help collect or report data which distinguish between students who read with understanding and those who mostly recognize or sound out words, and which also assess students' level of understanding.

Several tests are now available which can provide the data indicated above. Among these are the Degrees of Reading Power (DRP) developed by the College Board in order to overcome some of the obvious deficiencies of the Scholastic Aptitude Test and other similar tests. Another is the Word Test now being validated by Dr. Ron Carver of the University of Missouri-Kansas City. The DRP and/or the Word Test and similar tests are urgently needed to help ensure that instruction is not driven unproductively toward mastery of narrow word attack skills which do not add up to reading with understanding. This is particularly important for disadvantaged students, whose elementary and middle-grades reading scores probably improved in the 1970's, but who did not make adequate gains in reading with comprehension. Comparable tests are needed to provide improved assessment of problem-solving in math as well as other higher-order skills.

A related test which also should be used to track the educational system's progress in developing the most important skills—i.e. higher-order thinking skills—is the Lawson Formal Operations Test. This test can help not just in assessing gains in students' performance on higher-order skills, but indirectly can help determine whether science, social studies, and other subject areas are being taught to develop thinking skills rather than unproductive rote memorization.

As part of the analysis of students' performance regarding comprehension and thinking skills, the NCES should conduct an analysis in selected big cities of discrepancies in performance between scores on the tests indicated above, performance on standardized tests such as the ITBS, the Stanford Test, or the Metropolitan, and performance on state and local criterion-referenced tests. Although the conclusions of this analysis are predictable, i.e., many standardized tests and mastery tests currently in use frequently yield relatively high scores (particularly in the primary grades) when comprehension and other higher-order skills are low, and vice versa, documentation, verification, and publicity are urgently needed to avoid another decade of disaster in working to improve performance in big cities.

2. Higher order thinking skills beginning at 4th grade and going at least through Bachelors (particularly analyzing, synthesizing and evaluation) should become a common measure in our assessment programs.

Explanation: It is important to know at what age students begin to show real growth in this type of intellectual development. Additionally, the purpose of education ought to be in the final analysis to develop this type of intellectual development. There are tests available to do this, including the DRP.

3. Correlation between grades, achievement tests, and social class.

Explanation: It is clear that previous examination of how closely the grading system and testing program reflect social class has been inadequate. It is important to identify and examine these patterns. Also, it is important to identify if and when, or where, this relationship begins to change. These data are already available in schools; they need to be collected properly.

4. Studies of cohort groups nationally by social class in regard to attendance and achievement.

Explanation: Studies that show a relationship between attendance and achievement in schools could establish either a direct relationship or lack of a relationship, or determine that in some cases it matters and others it does not, i.e., are underclass children hurt by absence and upper middle class children not? Again, these data are available now but not correlated or collected in this manner. They could be.

5. Alienation from school for students beginning at grade 5 or 6 and through high school.

Explanation: Beginning as early as possible, data regarding alienation from school on the part of individual children or social class cohort ought to be collected. Such data when correlated with other data might reveal important information concerning context, climate, and learning. Instruments measuring this are available. There was a good deal of work on this in the 70's, both in the United States and Canada. Toronto University was particularly known for this.

6. The type and nature of communication that occurs between the school and the community.

Explanation: Data collected concerning the frequency, content, and type of communication between schools and homes of different social class and configuration (single parent, stepparents, foster houses, etc.), might begin to identify differences related to expectations. These data are difficult to collect and would require more time than some might want to spend in the collection process.

7. Information concerning achievement data by social class cohort relating to improvement from year to year.

Explanation: It might be helpful to have a better understanding of how much improvement takes place from year to year by social class. Does it change as higher grade levels are involved? How many students overcome social class disadvantage at differing grades? Is the improvement the same and social classes simply start at different levels? These data are available if appropriate collection procedures are used.

8. Holistic data regarding students writing improvement on a yearly basis.

Explanation: Qualitative data is difficult to collect nationally. However, performance data beginning in kindergarten concerning language development is crucial. Written essays could be collected in September and another set in May. (A national sample would be needed.) Teams to do holistic grading of the two essays would be able to identify improvement. These performance data would be highly reliable data for prediction of school success. They could be collected in the 100 largest school districts in the country.

9. Other data of importance not currently collected systematically are:

A. Statistical data regarding Teacher Absence and Student Suspensions with related reasons.

B. State and large district comparative statistical data which would be of value and interest include: Technology/Computer utilization and application; Extent and success of the high school reform movement; Success of the Effective Schools movement; Extent and success of teacher competency testing; Equity and high risk children (Desegregation, Teenage Pregnancy, School Dropouts, Bilingual students, achieving excellence with equity).

In regards to the equity and high risk children issue, the large school districts are concerned primarily with learning about what programs are being developed and implemented to help solve the problem.

C. The Digest of Education Statistics provides most all comparative data by states. There is one section where selected statistics for the 120 largest school districts are provided. The expansion of this particular section to include more teacher and student data would be useful.

Again, the issue is not how much and how frequently to collect data. It is whether we want to collect diagnostic data which will assist us in formulating policies which will do more than measure, that is give appropriate direction to initiate necessary changes required to improve performance on the part of educational institutions and students. The populations in our big cities do not require more and better assessment on the wrong issues; let us all direct our efforts toward collecting data in a manner to give guidance in an effort to improve performance.

An Elementary and Secondary School Statistics Program for the National Center for Education Statistics

W. Vance Grant
Chief, Statistical Information Office
National Center for Education Statistics

My goal in preparing this paper is to speak for the users of education statistics, those individuals and organizations that call, write, and visit the National Center for Education Statistics (NCES) in their search for meaningful data on elementary and secondary education in the United States. I know the kinds of questions they ask, and I think I know the kinds of statistical information that will be most useful to them.

My qualifications for speaking for the users are as follows: I arrived at the Center (it was known as the Research and Statistical Services Branch of the Office of Education in those days) in December 1955. For some months I worked on the surveys of City School Systems and State School Systems. Upon the completion of these surveys in the late summer of 1956, my work assignment was changed to include the statistical information function, and I have been closely identified with the Center's dissemination program for the past 29 years. During this time I estimate that I have talked with approximately 3,000 users a year, or a total of about 87,000 users during the 29-year period.

I have developed a list of key items in the field of elementary and secondary education that should be collected on a recurring basis. Except for those items designated as Census data, the items should be collected by the Center at the State level. The State figures should then be aggregated to provide national totals. Both State and national totals should be published in regularly recurring publications of the Center. Great care should be exercised to see that the figures are comparable from State to State and consistent from one year to the next. Consistent series of data enable us to measure trends over time, and this is of vital importance in our work. The list of key items follows.

List of Basic Statistics Frequently Requested from the Statistical Information Office

(The date of the latest published and/or readily available NCES statistics follows each item)

Public Elementary and Secondary Schools

Pupils
- Enrollment by grade* (Fall 1983)
- Enrollment by level (elementary vs. secondary)* (Fall 1978)
- Enrollment by age, race, and sex (Census data)
- Enrollments in high school subjects* (1981-82)
- Average daily attendance and average daily membership* (1980-81)
- Average length of school year and days attended per pupil enrolled (1980-81)
- Pupils transported at public expense (1980-81)

Employees
- Classroom teachers by level* (1980-81)
- Classroom teachers by sex* (1980-81)
- Classroom teachers by teaching field (1979-80)
- Other professional staff by type of position and by sex* (by type of position only, Fall 1981)
- Nonprofessional staff (Fall 1981)

Schools
- By level* (1982-83)
- By grade span (1982-83)

School districts
- By size of enrollment* (Fall 1981)
- Operating vs. nonoperating (Fall 1982)

High school graduates
- By sex* (1980-81)
- By type of program (Spring 1980 senior class)

Revenue receipts
- From Federal Government* (1982-83)
- From State governments* (1982-83)
- From local governments* (1982-83, including other sources)
- From other sources (gifts and tuition and transportation fees) (1967-68)

Nonrevenue receipts (1980-81)

Expenditures
- Current expenditures for regular school program* (1982-83)
  - Instruction* (1980-81)
    - Salaries of classroom teachers* (1981-82 estimates)
    - Salaries of other instructional staff* (1975-76 data for total instructional staff)
    - Salaries of nonprofessional staff (1975-76)
    - Free textbooks (1975-76)
    - School library books (1975-76)

- Supplies and other instructional expenses (1975-76)
- Administration* (1980-81)
- Operation and maintenance of plant* (1980-81)
- Fixed charges* (1980-81)
- Other school services* (1980-81)
  - Transportation of public school pupils (1980-81)
  - Health and attendance services (1980-81)
  - Food and other services (1980-81)
- Other current expenditures (summer schools, community services)* (1980-81)
- Capital outlay* (1980-81)
- Interest on school debt* (1980-81)

Private elementary and secondary schools

- Pupils
  - Enrollment by grade (Fall 1978)
  - Enrollment by level* (1970-71)
  - Enrollment by age, race, and sex (Census data)

- Employees
  - Classroom teachers by level* (1970-71)
  - Other professional staff (Requested in Fall 1978; not readily available)
  - Nonprofessional staff (Requested in Fall 1978; not readily available)

- Schools by level* (1980-81)

- High school graduates by sex* (1964-65)

---

*While all the items on this list are judged to be important, those marked with an asterisk are considered critical items if we are to continue to provide adequate service to the public.

All of the above statistics should be collected at least biennially, and some of the really basic items, including public school enrollment, attendance, teachers, graduates, revenues, and expenditures, should be collected on an annual basis. The annual figures should be published in the kind of report we used to call *Fall Statistics of Public Elementary and Secondary Day Schools*. In preparing this report, we should emphasize speed rather than precision, so that the data can be published before the end of the school year to which they relate. This means that the financial data in the fall report will necessarily be estimates rather than final, audited figures. When the fall survey is repeated, the respondents should be encouraged to report any changes that have occurred in the data they submitted for the previous year, and those corrections should be printed in at least one subsequent edition of the publication. As our model for this kind of reporting, we might very well look to the *Estimates of School Statistics*, published annually by the National Education Association.

In addition to the annual *Fall Statistics* report, the Center should also publish a definitive, comprehensive report on public elementary and secondary education. This report, which should be prepared biennially, will provide a detailed statistical account of public education in each State and in the Nation as a whole. It will contain all of the items on public schools listed above, and it may very well include additional information as well. It will provide an analysis of trends over time and will also devote considerable attention to interrelationships among the data items; e.g., enrollment will be compared with the number of teachers, and expenditures will be related to the number of pupils in average daily attendance. Our model for this report should be our own *Statistics of State School Systems*, which the Center published for many years but discontinued after 1975-76. A senior educational statistician with a thorough background in public school finance should be assigned the responsibility for this major study.

I consider the *Fall Statistics* and *State School Systems* to be the cornerstones of our elementary and secondary statistics program, and they deserve the highest priority when we are planning and conducting our surveys. I now turn to surveys of secondary or tertiary importance.

At intervals of two or three years, we should publish a directory of local public school systems. The directory, in addition to giving names and addresses, should provide a small amount of statistical information about each system. The following items should be adequate: enrollment, teachers, high school graduates, schools, and current expenditure per pupil. I believe that a directory could be designed that would provide all of this information in a publication about the same size as the one we published in the fall of 1980. The 1980 publication contained much less information, however. The directory should also contain a number of analytical tables that show the number of systems by State, by grade span, and by size of enrollment.

Every other year, for those years when we do not prepare a comprehensive *State School Systems* report, we should publish an abbreviated report on *Revenues and Expenditures for Public Elementary and Secondary Education*. Trends in public school finance are important, are in a state of flux, and should be measured annually. The publication should provide State and national totals on revenues by source and on expenditures by purpose but not necessarily in as much detail as the data in *State School Systems*.

We should conduct a survey of private elementary and secondary schools biennially. Private schools have increased in number and in enrollment, and they certainly deserve to be represented in our statistical program. The great need here is for a consistent series of State and national figures on schools, enrollment, teachers, and high school graduates. The data should be collected by affiliation of school, and the data on schools, enrollment, and teachers should be by level; i.e., there should be separate figures for elementary and for secondary schools. Nursery school children probably should not be counted in our data on elementary school enrollment. Most of these children are probably not involved in a truly "educational" program, and their inclusion in our statistics makes comparisons between public and private school enrollment almost meaningless. It was much more meaningful when we could compare public and private enrollment by grade groups (kindergarten through grade 8 and grades 9 through 12) or by level (elementary, excluding prekindergarten, and secondary).

One of our major studies that has been conducted rather infrequently is the survey of offerings and enrollments in high school subjects. While there are substantial difficulties inherent in a survey of this kind, the fact remains that these data are of great interest and value to the users of education statistics. It was certainly a breakthrough when we were able to obtain 1981-82 data from the survey of High School and Beyond, and I recommend that we request similar data from our respondents when we conduct longitudinal surveys in the future. Our eventual goal should be to obtain these data at intervals of about four to six years.

The survey of preprimary enrollment should be continued at intervals of two or three years. There is a good deal of interest in early childhood education, and this may intensify as the number of young children increases. This would appear to be the appropriate vehicle for measuring the participation of 3-, 4-, and 5-year-olds in nursery school programs. As indicated above, a problem arises when you include large numbers of these children in private elementary school enrollment along with small numbers of them in public elementary school enrollment and then compare total enrollment in public and in private schools.

Special education for the handicapped and for the gifted has been the focus of a number of studies of this office through the years. After a long period of inactivity in this area, we published a contract report, *The School-Age Handicapped*, earlier this year. The report used program data from the Office of Special Education and Rehabilitative Services, U.S. Department of Education, to provide national totals on the "number of children 3 to 21 served annually in educational programs for the handicapped." I would like to see the Center do a survey, possibly a large sample survey, in which we obtain State and national totals on the number of handicapped children and gifted children enrolled in special education programs. The handicapped children should be reported by type of handicap, and it would be interesting to know how many are being educated in regular public schools and how many in special schools for the handicapped.

A promising area for research is the organizational structure of public and private elementary and secondary schools. In the recent past we have seen the rise of the public middle school, a category consisting of schools with grades 6 through 8, 5 through 8, or some similar combination of grades. At the same time there has been some reduction in the number of public junior high schools, i.e., schools with grades 7 and 8 or 7 through 9. A study of these trends, the number of children affected, and the influence of different organizational patterns on the learning process would be of considerable interest.

Data from this office on the educational background, teaching assignment, years of teaching experience, and personal characteristics of public school teachers would be very useful. This kind of information is reported at five-year intervals by the National Education Association in their _Status of the American Public School Teacher_. Our survey, based upon data from a nationwide sample of public school teachers, should not be designed to supplant the NEA study, but it should provide more frequent data from a substantially larger sample.

Almost everyone is interested in the dropout problem, but no one seems to produce definitive data on the number and characteristics of dropouts and why they leave school before high school graduation. The logical place for a survey of this kind is the Longitudinal Studies Branch of the Center. This group has already followed up the high school sophomores of 1980 to see how many of them graduated in 1982. This survey should be regarded as a trial run, because it excluded those persons who dropped out of school before the spring of their sophomore year. When work begins on the next cohort of students, I recommend that the study measure dropouts from the beginning of the ninth grade. Currently, about 99 percent of the young people enter the ninth grade, but then the rate of withdrawal from school accelerates sharply, especially between grades 10 and 11 and grades 11 and 12. Most of the high school seniors do stay in school until graduation.

Another area that we talk about a great deal and where we have very little hard data is the quality of education. People want to be able to compare the education provided in their school system or their State with the education in other school systems or other States or with national norms. These comparisons are very difficult to make unless one is willing to settle for quantitative surrogates for quality, such as expenditure per pupil, average salary of classroom teachers, and pupil-teacher ratios. One solution to this problem is to use the national norms provided by the National Assessment of Educational Progress (NAEP) and to encourage the States and even some local school systems to administer the NAEP tests to their students. If enough States and communities participated in the program, we might eventually reach the point where we could begin to have some notion of the differences in the quality of education being provided in different areas. Obviously, the measures derived in this way would not be the final word, but they would be a beginning.

I don't want to leave the field of elementary and secondary education without saying a good word for our statistical projections program. This program is a legacy from our good friend and long-time branch chief, Dr. Kenneth A. Simon. The projections have appeared on a regularly recurring basis since 1964, and our statistical information staff uses them constantly. Especially in the past few years, when the flow of reliable information on elementary and secondary schools has slowed to a trickle, we have found the projections to be invaluable. We definitely should continue to produce annual projections at least 10 years into the future of school enrollment, teachers, teacher supply and demand, high school graduates, and expenditures for education at the national level. I should also like to see the program expanded to include State projections of enrollment, teachers, and graduates for the school year just beginning and for 5 and 10 years ahead.

Up to this point I have talked mainly about the content of our elementary and secondary program. This is appropriate because my work makes me data oriented rather than process oriented. But, in conclusion, I should like to say a little bit about methodology and sources of data.

In the mid-1970's there was a great deal of talk in the Center about a new survey, the Common Core of Data. This survey was designed to provide vast amounts of information from the local school systems around the country. About this time I had a foreign visitor, from India as I recall, to whom I was describing the Center and its data collection program. After I had done my best to describe the Common Core of Data, he responded with indisputable logic: "Why in the world would anybody go to 16,000 sources to get the same information he could get from 51 sources?" I thought he was right 10 years ago, and I still think so today. Nobody has ever been able to explain satisfactorily to me why we collect such a great mass of data on local school systems that are never published and that practically nobody ever sees. Instead we should be concentrating our limited resources on getting good trend data from each State department of education and on analyzing and publishing these State figures expeditiously. The data we publish on local school systems should be limited to the information in our directory of school systems plus a couple of tables in the _Digest of Education Statistics_ and/or _The Condition of Education_. And we certainly should not be collecting large amounts of information that we don't plan to use.

Our major source of statistics on elementary and secondary schools should be the State departments of education. For information that is not available in the State departments and that we still feel we urgently need, we usually should resort to sample surveys. Useful information can be obtained, for example, from the population surveys of the Bureau of the Census or from a nationwide sample of teachers. The survey of private elementary and secondary schools may very well be an exception. If we are going to do this survey at all, it should be done well. The data should be consistent from one year to the next and should be comparable with the figures we obtain for public schools. In addition, we need to be able to provide the private school data by State, by level, and by type of affiliation. In order to meet all of these criteria, a survey of the universe is probably required, even if it means that we have to conduct the survey less frequently.

If you have read between the lines of this paper, you have probably gotten my message. In case you haven't, the message is this: I found a home here, and I believe in the Center and its mission. On the other hand, even a good thing can be improved, and that includes the elementary and secondary statistics program of NCES. One of our major problems is not that we have done too little, but that we have tried, with limited resources, to do too much. We have spread ourselves so thin that the really important surveys, like Statistics of State School Systems and Fall Statistics of Public Elementary and Secondary Day Schools have fallen through the cracks. It is time to pick up the pieces and put them back together again.

Two Suggestions for NCES Data Collection

Jane Hannaway

Woodrow Wilson School of Public and International Affairs

Princeton University

Princeton, New Jersey 08544

609/ 452-5078

Paper invited by the National Center for Educational Statistics,
U.S. Office of Education, Washington, D.C. (Contract No. NIE-P-85-9539)

July, 1985

This paper suggests two areas for consideration as the federal data collection effort is redesigned. It is not intended to assess the overall federal effort, but rather to identify two specific policy questions which currently available data do not handle fully or appropriately. The first question is concerned with measures of process: How many and what type of personnel are employed by school districts and what do these personnel do? The second question is concerned with input measures. It is an old question: What are, and how do, various student input characteristics affect output, particularly student achievement? Output measures, per se, are not dicussed.

1. LEVELS OF FUNCTIONAL EFFORT

The general concern in the first question is with measures of functional effort in school districts. How many and what type of personnel are employed by school districts and what do these personnel do? The question is basic and deceptively simple; but getting accurate measures is difficult.

Of most concern here is the level of effort devoted to the administration of schools and school districts. There are two main parts of this discussion. The first part focuses on the importance of getting reasonably good measures of administrative effort; and the second part discusses how this might be done.

Indicators of administrative effort are useful measures in any type of organization. There is a natural tendency for the administrative part of an organization to grow and, because the marginal contribution of administration to productivity is difficult to ascertain, internal checks on growth are limited. In educational organizations, it is particularly important to get good measures of administrative effort. In addition to general problems of assessing the marginal contribution of administrative activity to the organization's welfare, there is the more specific problem in education of a production process that, itself, is not well understood. This compounds the problem of factoring out the extent to which administrative efforts, over and above other factors, contribute to productivity.

Although the technology of educational institutions is not well understood, these institutions still must be responsive to community needs. Educational institutions are important to society. In them reside the hopes of society for its future as well as the repository of the best of its past. Society has placed great trust in them and, not surprisingly, looks for verification of this trust. In general, educational institutions respond in symbolic ways rather than through actually modifying some central aspect of their process or their product. (Meyer and Rowan, 1977.) It is therefore possible for educational institutions to be very sensitive to consumer/community pressures, but to do so in ways that have very little effect on the actual productivity of the system. For example, simply assigning an individual to serve as a Director of Evaluation, or Community Relations, or Bilingual Education or Programs for the Gifted, can go a long way in satisfying demands for responsiveness to community concerns. Administrative responses are immediate and visible. Changing the "production process" is not only more difficult, but the effects of any changes on the system's output are uncertain and long term. To some extent this is functional for the organization. The central production tasks, i.e., the teaching and learning concerns, are buffered from the whims of the environment. But it also suggests that the contribution of administration to productivity may be less than what is commonly assumed.

A close watch should probably be kept on levels of administrative effort during times of reform and during periods of enrollment decline. The typical way in which almost all organizations respond to productivity problems is to improve management. In general, this makes sense. One of the important functions of management is to design production processes so that at least a minimum quality of work is performed. If the quality is lower than what it should be, it is management's job to do something about it. But if the link between administrative activity and production activity is not clear (Hannaway and Sproull, 1978), i.e., if it is unclear what administration should do, or is doing, to improve education, problems can emerge. Better management can simply become equated with more management, i.e., more supervisors, more rules, more requirements. Their immediate effect on satisfying external pressure for the system "to do something" may be great, but their longer term effect on educational productivity, i.e., student learning, may not be very powerful.

Unlike periods of enrollment growth, where increases in the size of administration may not be very costly for the organization (i.e., the proportionate expenditure on administration may not change), the cost of an increase in administrative size during decline could be quite high. Administration would increase relative to the other parts of the organization implying a reduction in real expenditures on direct service, i.e. student contact activities. Indeed, findings have suggested that during periods of enrollment decline the relative size of the administrative apparatus of school districts is greater than it is during growth (Hannan and Freeman, 1975; Hannaway, 1977; Freeman, Hannan and Hannaway, 1977).

If it were possible to measure the productivity trade-off between direct services (i.e., teaching) and administration, one might be able to calculate an optimal resource allocation scheme. But this is not possible. It seems reasonable to speculate, however, that beyond some base level of administrative support, the marginal productivity of a teacher is greater than that of an administrator. After all, a teacher impacts directly on student learning.

The growth of administrative systems, it should be stressed, is not necessarily due to the self-aggrandizement of administrators. No doubt many reformers truly believe that more administration leads to better education. And, indeed, some administrative practices may have a significant positive effect on educational productivity. Unfortunately, however, we do not have a good handle on the benefits of either different types of administrative activity or varying levels of administrative effort. Some information that we do have on administrative behavior at the central office, however, suggests that administrators prefer to engage in activities that relate to external agencies rather than in activities that relate to teaching and learning concerns in the system (Hannaway, 1985). This, of course, is very troublesome. It suggests that externally generated reforms could have significant effects on the volume of administrative activity and little effect on the conduct of teaching and learning activities. Our understanding of the relationship between administrati on and educational productivity, or more generally between different levels of personnel effort and productivity, can be improved by collecting and analyzing data that permit direct investigation of these issues.

Two different types of data are discussed below. Each represents a different level of data collection and each would provide information to similar questions but on different levels of specificity. Some of these questions are: What is the variation in the distribution of manpower effort across school districts and what determines different allocations? How do different distributions and levels of effort contribute to educational productivity. Sampling and cost considerations would vary according to the type of data needed and the specificity of the question.

a). Counts of Personnel in Different Personnel Categories.

This seems like fairly straightforward data; but personnel data can be categorized very differently. One way, for example, is to report the number of individuals in each district with different certificated status (e.g., the number of individuals certified as teachers, psychologists, administrators, etc.). But this can be very misleading for, at least, two reasons. First, individuals do not necessarily carry out functions defined by their certificated status. For example, an individual with an administrator credential could be teaching; or a teacher might be working on special project for the superintendent rather than in the classroom. Second, states use different definitions for the same category. What may be defined as an administrator is one state may be defined as a curriculum supervisor in another. Cross state comparisons, therefore, can be very misleading.

Personnel might be categorized in more useful ways. One possibility is to categorize personnel according to how directly they contribute to student learning. Those who spend a majority (or some other determined amount) of their time in direct interaction with students could be one category; those who contribute indirectly, e.g., principals and curriculum supervisors, could be a second category, and those who provide support for the system but who are not concerned with teaching and learning activities, e.g., personnel directors and accountants, might be considered a third category. It clearly would take some thinking to work out meaningful groupings, but the main point is that some accounting based on the extent of direct contribution to student learning could be useful and it should be done in a consistent way across states.

Some current personnel categories, such as 'other administrative' and 'other instructional' are quite large in some states and quite small in others (Digest, p.49-50). For example, in New York there are nearly 4 times as many 'other instructional' personnel than there are principals and in Florida there are more than 3 times; but, in Connecticut and Missouri there is not one person in this personnel category. Who are these 'other instructional' professional people and is it reasonable to expect that they make a direct contribution to student achievement? Is it possible that New York (and some other states) provides some additional and different instructional support for students that Connecticut and Missouri do not? If so, what is it and is it worth it? Or, do these figures, more simply, reflect the different way states certify and therefore categorize personnel?

Without good measures of the type suggested here, it is impossible to analyze how different reforms or different state administrative systems affect the allocation of personnel effort at the local level or to understand completely why reforms might (or might not) affect student achievement.

b). Time allocation of individuals. The personnel categories discussed above would give gross measures of functional effort across districts and states. Finer grained information, e.g., how much time administrators spend on different types of issues, could give a better picture of the types of work demands placed on educational professionals and how they respond to these. This type of data, however, is both difficult and expensive to collect. Information about districts under (a) could be collected regularly in a standard format across states; but the micro level data suggested here should be collected from a small sample of individuals/ districts/ states on only an occasional basis.

Examples of questions that this type of data can inform are: To what extent are principals focussed on teaching/learning concerns and to what extent on purely administrative chores? What are the administrative costs of categorical aid programs? How much time do principals/ counselors/ teachers spend working with parents? All these questions are concerned with how educational professionals themselves allocate their attention and effort. For instance, individuals serving as administrators of a special education program could be spending varying amounts of time talking to accountants, trying to keep track of the dollar flow; or to teachers, working out special classroom arrangements; or with parents, coordinating school and home support. One would expect that these activities contribute differentially to student learning and, from a policy perspective, we would want to structure the demands of the job in such a way that the greatest learning possible would take place. Information on task demands and how they are handled could help do this.

There are clearly some methodological problems in collecting time allocation data. (See Hannaway, forthcoming.) But it would probably be worthwhile trying. One way not to get estimates of time allocation is to ask respondents how they spend their time. At least in the case of administrators, this method is beset with biases. Managers are not very accurate reporters of what they do, probably because they engage in many varied tasks most of which last only a few minutes. The mental exercise they are asked to perform in making an estimate is complex; they must first recall and then aggregate thousands upon thousands of short tasks. Their estimates are based on recall and what managers are likely to remember are those tasks that are particularly vivid in their memory and those that fit with preconceived notions. Therefore, they tend to overestimate tasks they found particularly rewarding (or painful) and those tasks that conformed with their normative expectations of what someone in their position should do. When aggregating their tasks, they tend to underestimate tasks of short duration even though there may be many such tasks that together account for a large fraction of their time.

One way to proceed, which would be both reasonable methodologically and not terribly costly would be to use the diary method. This method has been used successfully with both managers (Stewart, 1967) and college presidents (Cohen and March, 1974). Either the respondent herself or her secretary would keep a log of daily activities. A selected number of individuals representing a personnel category, say, special education administrators , might keep track of what they do for one day. The results from a reasonable sample for any particular day could be quite informative. Similar studies of the time allocation of other types of personnel, e.g., principals, could also be done. This type of study would require more thought, organization and coordination than a standard survey, but the results might be well worth the effort.

Both of the above suggestions for data are based on the simple assumption that in order to get a better understanding of the determinants of educational output we should get a better understanding of educational process. There are researchers who are asking similar questions at the classroom level which, no doubt, will be fruitful. I am suggesting that information about the direction and intensity of functional effort across the district as a whole could be valuable itself as well as a complement to classroom level information.

2. PUBLIC AND PRIVATE SCHOOL INPUT

There has been a considerable amount of attention and debate given to the relative success of public and private schools in producing cognitive achievement; i.e., the output of the two systems. Much of the debate and controversy has centered on the input. Critics argue that private schools do better because they are working with different types of students and that standard background measures, such as those used by Coleman, Hoffer and Kilgore, do not capture these differences. This is an important public policy debate and one that current data do not allow us to sort out completely.

The only way to get accurate measures of the effect of private schools (or different types of public schools) is to factor out in a completely reliable way self°2Dselection effects. This can be done with some confidence statistically; but collecting parent/ student measures prior to schooling choices would be preferable. The basic questions is: Are private school students (and parents) different from their public school counterparts before they even enter private schools? For example, does private school selection, by itself, indicate higher levels of parent and student motivation and commitment to education; or do private schools, and, perhaps, the very act of choice by parents foster certain attitudes, values and behaviors?

To truly distinguish self-selection effects from school effects requires a different type of data collection effort where, perhaps, a small number of communities are selected and the attitudes, experiences and choice behavior of members of those communities, i.e., parents and students, are studied over time. From this, it would be possible to estimate the extent to which public and private school parents/students are different as well as the experiences that contribute to parents opting out of (or staying in) a particular school or school system.

Such an effort should not be considered a substitute for data already being collected on private and public school comparisons, but rather an additional effort to address a very specific and very basic question about the characteristics of public/private school input.

Input differences may also be important for teachers. Are individuals with different characteristics and values attracted to public and private school teaching? How does the culture of the school affect teacher behavior and attitudes. Answering these questions would also best be done using some type of longitudinal data collection design.

While the returns from such an effort would probably not be immediate, studying teachers in this way could be quite profitable.

It has long beer recognized that professionals are affected by normative expectations communicated during their training and on the job, especially their first job. And it is not unreasonable to expect that the normative structure of a school is affected by its institutional arrangements. Consider, for example, the different roles that parents play in different types of schools and the likelihood that parents convey expectations and rewards different from other involved parties, e.g., state bureaucrats or union representatives. If information were collected from teachers over time, say, in a community study, it could lead to a better understanding of the determinants of teacher behavior. That is, it would be possible to track teacher characteristics/ behaviors and estimate the extent to which these are affected by institutional arrangements and parent/student characteristics.

This paper focussed on two weaknesses in the current data collection effort of the federal government. One was concerned with a process measure and the other with an input measure. These weaknesses limit our ability to address policy questions about the educational process and how it is affected by various reforms and institutional arrangements. There was no discussion of output measures. The responsibility of the federal government in this regard is less clear. While some overall assessment of the state of educational productivity in the country is within its purview, education reforms are being formulated and administered at the state and local levels. And it is information at the school level that will be most useful in evaluating and refining the reform effort. This level of information is probably best collected by states and districts who will also be the users.

REFERENCES

Cohen, Michael and James G. March

1974 Leadership and Ambiguity. New York: McGraw-Hill.

Freeman, John, Michael Hannan and Jane Hannaway

1978 "The Dynamics of School District Administrative Intensity: Effects of Enrollment and Finances in Five States". Report submitted to ASPE, DHEW.

Hannaway, Jane

1977 "The Effect of External Funding on the Administrative Intensity of School Districts." Paper presented at the Annual Meetings of the American Education Research Association, New York.

Hannaway, Jane

forthcoming "Managerial Behavior, Hierarchy and Uncertainty: Notes Toward a Synthesis," Human Relations.

Hannaway, Jane and Lee S. Sproull

1978 "Who's Running the Show: Coordination and Control in Educational Organizations". Administrator's Notebook, Chicago: Midwest Administration Center, University of Chicago, 27:9 (1979).

Meyer, John and Brian Rowan

1977 "Formal Structure of Organizations as Myth and Ceremony," American Journal of Sociology, 83: 340-63.

Stewart, Rosemary

1967 Managers and Their Jobs. London: Macmillan.

THE QUEST FOR EXCELLENCE/PUPIL SELF-ESTEEM

Dr. Carole Hall Hardeman, President
ADROIT Publishing, Inc.
Oklahoma City, Oklahoma

June 19, 1985

A Paper Commissioned by the U.S. Department of Education, Division of Research and Statistics; under auspices of the
National Center for Educational Statistics
(This paper may not be copied without permission of NCES)

ABSTRACT

This paper examines concerns about the educational process in relationship to policy issues and administrative needs and measurement of our Nation's education systems.

This paper identifies specific data elements that can provide necessary information in support of present and future government, business, and academic decisionmaking, and that can help inform the American public.

Specific issues addressed in this paper include the following:

- Opening statement of self-worth and its importance in the search for excellence
- A discussion centered around the problem of how self esteem can be developed in the classroom
- Areas for new data bases which include:
  Desirable qualities of teachers and their impact on student achievement;

  Which of these qualities are essential for measurable academic achievement;

  Teacher qualities essential for the development and enhancement of pupil self-esteem and pupil creativity;

  Methods of evaluation for non-traditional teaching methods;

  Weighing out affective attributes of teachers;

  Improving "state of the art" by correlating effective student outcomes with well-defined successful teacher behaviors and competencies.
- Description of "Teachers I Have Known"
- Teacher ability to define classroom objectives and to facilitate excitement/motivation in the classroom

THE QUEST FOR EXCELLENCE/PUPIL SELF-ESTEEM

Introduction

> *Probably the most important reason for striving for excellence is because it creates opportunity. ....It goes without saying, that it is the students who have achieved excellence who will have the best opportunities for challenging work, top pay, and access to graduate school.......Another reason to strive for excellence is because it is basic to your self esteem.......We all have a considerable investment in our education and we spend one third of our working lives on the job. WE HAVE A NEED FOR SELF-WORTH AND THAT COMES FROM DOING A JOB AS BEST WE CAN.*
>
> R. B. Powell

The above quote(s), taken from an address at the annual conference of the Mexican-American Engineering Society (MAES) can and do apply to students or teachers of any race, and to all age groups. Powell stresses the idea that although we are usually rewarded for excellence in schools or on the job, there is also a feeling of self-satisfaction that comes from doing a job well. This sense of pride cannot be taken away and it cannot be diminished by the fact that there may be no specific monetary reward attached to a given assignment. "Self-esteem", he continues, "is important, and fortunately, it is a by-product in the search for excellence".

For the past six years, while developing a science curriculum program which emphasizes successful scientist role models who have overcome numerous obstacles, I have been extremely impressed by the healthy self-esteem of over one hundred of the SOUNDS OF SCIENCE role models.

Since these materials are designed primarily for use in the upper elementary and middle school grades, great efforts have been made to relate to pre-adolescents the lifestyles, likes and dislikes, hobbies, favorite school subjects, personality traits, and other useful data about these scientist role models when they were pre-adolescents. The majority of these role models did not have extremely high self-esteem as youngsters. Some of their self-described traits as youngsters are loneliness, self-appraisals of being physically unattractive, experiencing aversion to certain academic subjects, having little knowledge of what they would become as adults, etc. The great majority of these persons, however, can still remember a significant adult who encouraged them to be the best of whatever they pursued, and at least fifty percent of these scientists were encouraged to study mathematics and science in high school. Few respondents can remember being encouraged by their junior high or high school teachers to pursue careers in science. Noting the difficulties of generalizing and thus diminishing the results of two years of gathering data on over 350 successful scientists, the following characteristics are reported only because they accurately describe 90% of the scientists: They loved pets and spent a great deal of time with their individual pets; they had specific tasks to perform around the house or had after-school jobs for which they received earnings; and most significantly, they belonged to several extra-curricular clubs, sports organizations, music organizations, (band, choir,) and they were always busy. How much did these activities contribute to their eventual success as scientists?

> *A final reason why striving for excellence is so important to minority engineers is because it serves as an example to fellow students, co-workers, and to the community, which others can follow. This is especially important to various Minority groups where there is a shortage of role models in many fields, including science and engineering. And by having more role models who are striving for excellence in school or on the job, the young students will also respond and follow that lead because they will know that striving for excellence is desirable and achievable.*

The question to be addressed in this paper is centered around the problem of how self-esteem can be enhanced in the classroom. The aforementioned scientists did not have high self-esteem as youngsters, according to their own individual assessments. How did they achieve the high level of success which they now enjoy and at what point did self-esteem develop or evolve, and more importantly, what set of conditions must be present in the classroom to facilitate or bring about a student's sense of great self-worth?

I have chosen to not dwell on the all important areas of cultural pluralism, SES of pupils, parental involvement with student learning, etc., because the data on these topics has been more than adequately documented. Instead I have chosen to limit this discussion to the qualities of teachers and the quality of the day-by-day experiences in the classroom. There have been numerous papers written and countless workshops and seminars and or conferences devoted to the causes of student failures, and all too often, the blame has been placed on the lack of student readiness to learn, to parental apathy, cultural and economic plights, etc. It seems to me that only in education do we place so much emphasis on what happened yesterday, last year, and other historical events as we absolve our own inability to solve the problem of what we, the professional educators can do to help millions of students to achieve academic success. Does the mechanic ask the owner of the automobile how it happened that the brake lining wore thin? When approached to change a flat tire, does the service station attendant ask "Where were you when you ran over this nail?" Is there one set of instructions for changing the oil if the car owner has just returned from a 5,000 mile trip as opposed to the negligent driver who drove the car around town for a year and forgot to have the oil changed? If a car window is broken by a burglar, is there a different method of installing a new window than if the car were involved in a collision? Does the dentist use a different method to remove a tooth if the patient broke the tooth in a fight, as opposed to breaking the tooth while accidently biting down on an unknown object?

The acknowledgment of cause of breakdown may indeed impact on the method of repair, but the quality of the finished product remains unaffected. The point is that teachers, educators, administrators,..our education profession in its totality; must cease and desist the practice of concentrating on the various causes for our pupils' failure to achieve and move to affectuate quality classroom experiences which will insure student excitement, inquisitiveness, and a genuine desire for knowledge. Instead of shaking our heads in disgust and occasionally with contempt for those students who have difficulty making their way through the educational maze, we must set into motion a set of conditions which will result in a well-ordered, self-disciplined, orderly school environment without which learning cannot take place.

In many disciplines, research reports are based on experimental data. The mere nature of education of youngsters dictates that our research will be empirical in nature, and most of educational theory is therefore based on this type research. Empirical data is most useful when we set out to study entities designed to improve our own impact on society, for future generations:

- What are those qualities of teachers which are positive and desirable in an effective school?
- What are those qualities, though positive and desirable, which have little or no impact on student achievement?
- How many of these qualities or variables carry heavy weight in the hiring process? In the evaluation process?
- Which teacher qualities(prioritized) are essential for measurable academic achievement of students?
- Which teacher qualities are essential for the development of high pupil self-esteem, for the enhancement of pupil creativity, and can these teacher qualities be measured by traditional methods?
- What methods can be utilized to evaluate essential qualities which cannot be measured by traditional methods?
- How much weight should be attached to the following teacher qualities?

  Great personality
  Warmth, caring attitude
  Attractive physically
  Ability to "get along" with other faculty and administration

- Are there consistent specifically identifiable characteristics unique to the students of certain sucessful teachers in a given school? Can these characteristics be isolated and correlated with the learning environment and process(es) from which these students benefitted?

  Could such information be beneficial to the "state of the art"?

TEACHERS I HAVE KNOWN

I have known some very unpopular teachers; scorned by their colleagues, and highly respected by their students. These teachers were known to accept nothing but the very best from all of their pupils. They had no patience with student mediocrity and they spent no time with borderline students. They pushed the "real scholars" almost beyond their limits. They challenged their bright students,..excited them,..leaving the "dummies" by the wayside. IS there a place for teachers such as this in the American educational system?

I have known sympathetic teachers who could get unbelievable results from "slow pupils",--make them feel terrific, smart, useful, indefatigable in their quest for more knowledge, for the perfect score! These teachers were criticized often by the "smart high IQ students" who became impatient with the repetition, the praise (who needs it?), the compassion. Is there a place for teachers such as this in the American educational system?

I've known teachers who couldn't get to school in time to "check-in" by 8:00 AM but who managed to always be in place by 8:15 for their First Hour Class which began at 8:30 AM. They usually stayed in their classrooms long after the 3:30 PM "sign-out"; maybe until very late in the evening. They were "night people" who blossomed late in the day. They stayed until they were satisfied with the results which they diligently sought. Since time was not that important to them..they considered it not at all unusual to work after school, or on weekends with a student who needed the extra time, or an instrumental or vocal ensemble who needed extra rehearsals to ensure that the performance met the highest standard. These teachers were motivated by one entity; excellence! Further handicapped by their inability to get required reports turned in on time, these teachers' inability or unwillingness to submit timely reports placed undue inconvenience and aggravated baggage on the principal or on the principal's secretary. These unorthodox teachers almost always caused the school showcase to become overburdened by the countless Superior trophies garnered by their students in Debate, Gymnastics, Athletics, Science Fairs, Academic Olympics, Vocal and Instrumental Contests, etc. But, their report cards were never turned in on time, they were late for faculty meetings, and sometimes they even had the unmitigated gall to MISS IMPORTANT FACULTY MEETINGS! (Coaches work is important, but the line must be drawn somewhere!) Is there a place for teachers like this in the American educational system?

Bloom contends that teachers are seldom hired for those qualities which are essential to effective pupil achievement. I contend that too little information has been documented on the subject of teacher qualities which are measurable and/or widely accepted from school system to school system.

TEACHER/LEARNER OBJECTIVES DEFINED

The final area of this paper addresses a concern which I feel is perhaps discussed in Teacher Education 101, but which is seldom utilized by 99% of today's teachers. How many teachers ask themselves the question: What do my students know at the end of this 55 minute class period that they did not know at the beginning of the hour? What is the objective of this lesson today? What are the entry level cognitive skills needed by my pupils in order to understand today's lesson?

How many teachers actually set aside the last portion of each class period to review the last 45 minutes? How many teachers tell the students what the next class period will involve? Can the students relate each lesson with a part of their own world?

In conclusion, I return to the original question regarding the enhancement of self-esteem: What set of conditions must be present in the classroom to facilitate or bring about a student's sense of great self-worth"?

Does the achieving of excellence and the knowledge that one has successfully completed an assignment increase self-worth? Does success breed success again and again? Do teacher qualities need to be redefined and must teachers become a part of the process? Do role models enhance a student's desire to strive for excellence because they observe that excellence is achievable?

Is there need for diversity within the teaching ranks in a given school? Should teachers be matched with the type pupils who seem to thrive under their tutelage? Are the "personality types" rewarded for getting along with their colleagues and conversely, are the introverted teachers' effectiveness in the classroom often unrewarded? Are teachers evaluated by persons who are guided by a set of criteria that is just and fair and more importantly, is the evaluation process designed to enhance the educational environment?

Of course, it's easy to answer these questions, but is it easy to set up a system by which they can be effectively measured, reported, and disseminated to the classroom teacher and the administrator in those remote communities who do not encourage attendance at conferences where such data is reported? Is it now known the percentage of educators who take advantage of educational statistics and reports such as those published by NCES?

I intentionally avoided a long dissertation on the issues of science and mathematics education and the serious lack of knowledge of the usefulness of these academic areas to the present and future lives of today's students. It seemed self-serving; as I am a publisher of math and science curriculum materials. However; I have taken the liberty to include as an appendix, some new information recently distributed by Howard Adams, of GEM. (See reference page for additional information of Dr. Adams). It seemed vital to share this information as part of this report. Adams' complete report further enhances the material included in my introduction about the relationship between role models, academic success, and career aspirations particularly as this issue relates to minority students. Problems faced in graduate and undergraduate school must be addressed in elementary and secondary school.

The attempt has been made to raise questions for this commendable study undertaken by the National Center for Education Statistics. The following suggestions are made in response to QUESTIONS TO BE ADDRESSED from Attachment A:

1. In answer to question #1, please refer to page 5 of Attachment B

Public School Survey:	ADD:
Summary level	Characteristics principals seek in hiring new teachers
Periodicity	Characteristics principals seek in evaluation of teachers
Data set (1984-85	Characteristics teachers believe are important in hiring/evaluating of teachers
	Characteristics needed for competent effective teachers as described by parents, community, students
	What has been done to upgrade Teacher Staff Development since 1980(as perceived by teachers?

2. In answer to question #2, please refer to page (un-numbered 9) CURRENT ELEMENTARY/SECONDARY EDUCATION DATA ACQUISITIONS

Under Fall Membership
  Public by:
  Private by:

Percent minority

The suggestion is made to specify races. Black children are quite different from Asian Americans. Asian Americans should be delineated by number of years in USA. Asian Americans who have been in America for ten years have had a totally different experience from those newly arrived Asians.

Many economically low Asian-Americans surpass Black Americans academically. Their lifestyles and American experiences should be chronicled for the BENEFIT of many Blacks who could perhaps appreciate "How do they do it?"

Hispanics and American Indians are also different from Blacks and Asians. The point: SPECIFY RACE.

NOTES

SOUNDS OF SCIENCE is a science curriculum program developed by Carole Hardeman et. al. This program is becoming widely used in several American school systems. It is unique in that it features on audio-cassette the lifestyles, childhood experiences, and career information of successful scientists of all races, persons with physical challenges, and over fifty percent of the role models are women. Although career based, this program engages students in excellent science laboratory activities which were developed by Otis Lawrence, Ph.D., Ray Broekel, Lh.D., and Richard Baim. Developed at the University of Oklahoma, the program is now under copyright to ADROIT Publishing, Inc.

REFERENCES

POWELL, R.B.,(1984). Why is Striving for Excellence Important? The Minority Engineer/Winter 1984-85, 73-74

BLOOM, BENJAMIN, All Our Children Learning. McGraw-Hill. 1981.

APPENDIX

ADAMS, HOWARD G., Engineering Education for Minority Students: A Status Report. (Dr. Adams is Executive Director of National Consortium for Graduate Degrees for Minorities in Engineering, Inc. (GEM)

Excerpts from his report are included as an appendix to this paper as informational data which might be useful. Additionally, it lends additional weight to some parts of the introduction to this paper.

APPENDIX

TABLE II

FULL-TIME FRESHMAN ENROLLMENT IN ENGINEERING BY ETHNICITY, 1973-1983

YEAR	TOTAL	TOTAL MINORITIES		BLACK		HISPANIC		INDIAN	
		NO.	%	No.	%	No.	%	No.	%
1973	51,920	2,987	5.8	2,130	4.1	790	1.5	67	.1
1974	63,440	4,018	6.3	2,848	4.5	1,068	1.7	102	.2
1975	75,343	5,344	7.1	3,840	5.1	1,384	1.8	120	.2
1976	82,250	6,315	7.7	4,372	5.3	1,766	2.2	177	.2
1977	88,780	7,133	8.0	4,728	5.3	2,161	2.4	244	.3
1978	95,805	8,792	8.7	5,493	5.7	2,662	2.8	225	.2
1979	103,724	9,792	9.4	6,339	6.1	3,136	3.0	317	.3
1980	110,149	10,399	9.4	6,661	6.1	3,373	3.8	365	.3
1981	115,280	11,116	9.6	7,015	6.1	3,689	4.1	412	.4
1982	115,303	10,721	9.3	6,715	5.8	3,633	3.2	371	.3
1983	109,638	11,478	0.5	6,342	5.9	4,760	4.3	376	.3

Source: Annual Reports, Engineering Manpwer Commission.

PARITY: Although the proportion of first year minority students entering engineering has risen significantly since 1973, the gains have done little to bring about parity within engineering education. When "The Effort" began, underrepresented minorities comprised 16.0 percent of the population and 5.8 percent of first year engineering enrollment (a difference of 10.2 percentage points). U.S. Census data for 1983 shows that the minority population had increased to 19.5 percent. Engineering first year enrollment for minorities in 1983 stood at 10.5 percent. The percentage difference for 1983 between population and first year enrollment was 9.0 percent.

TABLE III

POPULATION STATISTICS AND ENROLLMENT FOR MINORITIES IN ENGINEERING

	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983
Population Statistics (In Thousands)											
Total Population	209,936	212,227	214,542	216,883	219,250	221,642	224,060	226,505	228,976	231,474	234,000
Minority Population	34,769	35,735	36,728	37,748	38,797	39,875	40,983	42,122	43,292	44,495	45,731
Percent Minority	16.6	16.8	17.1	17.4	17.7	18.0	18.3	18.6	18.9	19.2	19.5
Minorities As a Percentage of First Year and Total Enrollment											
First Year Enrollment	5.8	6.3	7.1	7.7	8.0	8.7	9.4	9.4	9.6	9.3	10.5
Total Enrollment	4.6	5.2	5.5	6.0	6.1	6.7	7.0	7.9	8.3	8.1	8.4
Minority Recipients As a Percentage of Total Degrees Awarded											
B.S.	2.9	3.5	3.9	3.9	4.0	3.7	3.7	4.1	4.4	4.5	4.8
M.S.	1.5	2.2	2.0	2.1	2.2	2.7	2.3	2.4	2.6	2.3	2.9
Ph.D.	0.7	0.9	1.5	1.2	1.4	1.7	1.5	1.6	1.4	1.4	2.0

Source: Annual Report- Engineering Manpower Commission and U.S. Department of Commerce, Bureau of the Census

To reach parity with population percentages, all minority constituent groups will need to experience increased enrollment rates. Using 1983 Census data (Table III) for parity comparisons, minority enrollment should be 19.5 percent in engineering.

TABLE IV
FULL-TIME UNDERGRADUATE ENGINEERING ENROLLMENT BY ETHNICITY, 1973-1983

YEAR	TOTAL NO.	TOTAL MINORITIES		BLACK		HISPANIC		INDIAN	
		NO.	%	NO.	%	NO.	%	NO.	%
1973	186,700	8,558	4.6	5,508	3.0	2,769	1.5	281	.1
1974	201,100	10,530	5.2	6,287	3.4	3,380	1.7	323	.2
1975	231,379	12,828	5.5	8,389	3.6	4,111	1.8	328	.1
1976	257,835	15,412	6.0	9,818	3.8	5,138	2.0	446	.2
1977	289,248	17,753	6.1	11,388	3.9	5,747	2.0	618	.2
1978	311,237	20,729	6.7	12,954	4.1	7,150	2.3	625	.2
1979	340,488	23,999	7.0	14,786	4.3	8,454	2.4	759	.2
1980	365,117	28,944	7.9	16,181	4.3	11,860	3.3	903	.2
1981	387,577	32,196	8.3	17,611	4.5	13,615	3.5	970	.3
1982	403,390	32,711	8.1	17,598	4.4	14,035	3.5	1,078	.3
1983	406,144	34,126	8.4	17,817	4.4	15,182	3.7	1,127	.3

Source: Annual Reports: Engineering Manpower Commission

Trends in Full-Time Total Undergraduate Engineering Enrollment By Ethnicity, 1973-1983

Total enrollment for all engineering students increased from 186,700 in 1973 to 406, 144 in 1983. During the same period, total minority engineering enrollment rose from

1973 to 3,817 in 1984 (Table V). As a percentage of total B.S. degrees awarded in engineering, the increase was from 2.9 percent in 1973 to 5.0 percent in 1984.

The distribution of B.S. degrees among minority groups in 1973 was: Blacks 1.5 percent; Hispanics 1.3 percent; and American Indian 0.15 percent.

Advanced Degree Graduates

The national effort to address the underrepresentation of minority students at the graduate level was launched with the founding of the National Consortium for Graduate Degrees for Minorities in Engineering, Inc. (GEM) in 1976.

TABLE V

B.S. Engineering Graduates in The United States by Ethnicity, 1973-1984

Year	Total B.S.	Total Minorities		Black		Hispanic		Indian	
		No.	%	No.	%	No.	%	No.	%
1973	43,429	1,255	2.9	657	1.5	566	1.3	32	.07
1974	41,010	1,423	3.5	756	1.8	636	1.6	31	.07
1977	39,718	1,582	4.0	844	2.1	702	1.8	36	.09
1978	45,753	1,679	3.7	894	2.0	748	1.6	37	.08
1979	52,161	1,943	3.7	1,076	2.1	808	1.6	59	.11
1980	58,413	2,383	4.1	1,320	2.3	1,003	1.7	60	.10
1981	62,615	2,728	4.4	1,445	2.3	1,193	1.9	90	.14
1982	66,652	3,007	4.5	1,646	2.5	1,270	1.9	91	.14
1983	72,122	3,493	4.8	1,862	2.6	1,534	2.1	97	.13
1984	76,576	3,817	5.0	2,022	2.6	1,683	2.2	112	.15

Source: Annual Reports, Engineering Manpower Commission.

TABLE VI

M.S. ENGINEERING GRADUATES IN THE UNITED STATES BY ETHNICITY 1973-1984

Year	Total M.S.	Total Minorities		Blacks		Hispanic		American Indian	
		No.	&	No.	&	No.	%	No.	%
1973	17,152	258	1.5	104	.6	139	.8	15	.08
1974	15,885	345	2.2	158	1.0	187	1.2	4	.03
1975	15,773	320	2.0	141	.9	176	1.1	3	.02
1976	16,506	351	2.1	154	.9	183	1.1	14	.08
1977	16,551	364	2.2	147	.9	210	1.3	7	.04
1978	16,182	439	2.7	201	1.2	234	1.5	4	.03
1979	16,036	366	2.3	152	1.0	205	1.3	9	.02
1980	17,229	415	2.4	162	.9	249	1.5	4	.02
1981	17,643	466	2.6	182	1.0	276	1.6	8	.05
1982	18,289	414	2.3	184	1.0	215	1.2	15	.08
1983	19,673	580	2.9	258	1.3	306	1.6	16	.08
1984	20,992	636	3.0	253	1.2	358	1.7	25	.12

TABLE VII

Ph.D. ENGINEERING GRADUATES IN THE UNITED STATES BY ETHNICITY 1973-1984

Year	Total M.S.	Total Minorities		Blacks		Hispanic		American Indian	
		No.	&	No.	%	No.	%	No.	%
1973	3,587	26	.7	13	.3	12	.3	1	.03
1974	3,362	31	.9	12	.4	19	.6	0	.00
1975	3,138	47	1.5	17	.5	28	.9	2	.06
1976	2,977	35	1.2	10	.3	15	.5	0	.00
1977	2,814	39	1.4	16	.6	22	.8	1	.04
1978	2,573	43	1.7	15	.6	25	1.0	3	.10
1979	2,815	41	1.5	19	.7	22	.8	0	.00
1980	2,753	45	1.6	19	.7	25	.9	1	.04
1981	2,841	39	1.4	16	.6	20	.7	3	.10
1982	2,644	38	1.4	20	.8	15	.6	3	.10
1983	3,023	60	2.0	19	.6	41	1.4	0	.00
1984	3,234	49	1.5	24	.7	25	.8	0	.00

Source: Annual Reports, Engineering Manpower Commission

Obstacles to Success and Future Concerns

Underrepresentation of minorities in engineering is primarily caused by two obstacles:

1. Lack of knowledge about engineering as a profession; and,

2. Lack of preparation at the precollege level in math and science.

For the first area, the sparcity of professional engineers within the minority community leaves a void of role models. This results in minority youth not viewing engineering as a viable career path that would lead to success.

The second area is caused by the problems minorities encounter moving through the educational system in this country.

Some of these problem areas are:

1. Attrition from the pipeline--only 72 percent of Black and 55 percent of the Hispanic and American Indian students complete high school (Figure I).

2. Curriculum placement--only 28 percent of minority students pursue academic programs in high school. The rest are in either vocational (32%) or general (40%) programs.

3. Course Choice--among minority high school students only:

   50% take Algebra I
   35% take Geometry
   30% take Algebra II
   15% take Trigonometry
   5% take Calculus
   25% take Chemistry
   20% take Physics

The complexity of the problem is exacerbated by the manner in which the academic abilities of minority students are viewed. For example:

1. The educational system continues to use negative counseling when advising minority students. In the counseling process, minority students are too frequently counseled away from college preparatory programs and into nonacademic areas. Thus, minority students are often counseled below their abilities.

2. Some teachers fail to challenge minority students by teaching down to what they perceive to be the level of the student. In their efforts to offer under-

standing, they actually hinder academic growth by telling the student, "I know you are doing the best that you can do."

3. School personnel charged with counseling and advising minority students, too often associate learning difficulties with the economically disadvantaged. They tend to think that if a child is minority, he/she is automatically deprived and therefore probably cannot function educationally.

Because of these and other negative attitudes about the academic abilities of minority youth, the educational system,in spite of all the talk about change, remains ineffective in providing these students with a sound educational grounding in academic subjects. Thus, 75% or more of all minority students who graduate from high school do so unprepared to gain regular admission to post-secondary institutions of higher learning. That is why the work that is being done through the "Minority Engineering Effort" is so important and needs to be strengthened, expanded and continued. Future consideration needs to be focused on finding ways to:

1. Continue support for public education which is paid for by taxes. The notion of a voucher system of education will not work for the masses. If the U.S. is to have an educated populous, it will have to be done through a public system of education.

2. Identify academically talented minority students early in their education and counsel them into both academic curricula and courses that contain prerequisite preparation for college level work in engineering and/or other math-based disciplines.

3. Have teachers and administrators, at all levels of the educational structure, become positive about the abilities of minority students. In so doing, they must be made to realize that neither talent nor ability is defined by ethnicity. They must be made cognizant of the fact that students don't start out as failures--rather, the system designates who shall succeed and who shall fail. Through attitudes and expectations, the system conveys to minority students what is perceived as their chances for success. Studies show (Astin 1982, Brooks, 1983) that where the level of expectation is high and is com-

Review of Elementary/Secondary School Data
Needs of the National Center for Education Statistics

June 18, 1985

Forrest W. Harrison
U.S. Department of Education, Retired
formerly
Chief, Statistical Information Branch
National Center for Education Statistics

Introduction

This review of the elementary/secondary school data program of the National Center for Education Statistics (NCES) was approached by this author from a professional background which included positions as teacher, school administrator, educational statistics researcher, and manager of the statistical analysis and education statistics information branch of the National Center for Education Statistics (NCES). As I did then, I believe that NCES must serve the school administrators (State, intermediate, and local), the Department of Education, and the inquiring public, which includes the ordinary taxpayer, the news media, and the research community.

The collection, analysis, and dissemination of needed data must be done with an eye to keeping the task manageable in Washington, and the burden on respondents as small as possible while producing data which have reliability and validity. All these purposes can be served by increased dependence on sample surveys directed by knowledgeable professional educational statisticians and with increased cooperation of the State Departments of Education (SDOE) and the U.S. Department of Education. The problems of comparability of data between States and reliability of final products would be diminished by greater use of sampling.

For example, the quagmire of average salaries of teachers could be managed if an educational statistican, who understood State school finance, were in charge of a sample survey of about 5,000 local education agencies (LEA's). The sample size necessary may be more, or less, but it should be capable of producing a distribution of average salaries of teachers by size of district and by State. Where there are unusual arrangements, such as State payments directly into the teacher retirement funds, a small research project would be presented. However, if the researchers were working with a small number of survey forms, that problem could be solved by a system of attribution, which would make the figures comparable to the typical State figures.

Similar situations, such as State payments for construction of buildings and payments for debt service, could be managed by attribution techniques to make revenues and expenditures comparable to the large majority of independent LEA's.

Data Items or Data Series Needed

The 1985 data collection program as presented to the reviewers is impressive on paper and most of the work seems desirable; however, there are items and series for which there may be little need. The items most needed are those that describe schools, school districts or other local education agencies, pupils (including information about completers, dropouts, and graduates), revenue or income, expenditures, and outcomes. Much of this information can be had from records in the State capitals.

Schools and School Districts. Universe maintenance as described in the 1985 NCES program should provide sufficient data on schools and school districts. While it is essential to have these universe lists and to keep them up-to-date, the data there must be readily available to Department employees and other researchers on a timely basis and provisions must be made to provide information on a purchased-service basis as needed by the public. The maintenance of these files would be further justified if directories were produced on some reasonable cycle. It has been about seventeen years since school directories were published by the Department and LEA directories cannot be counted on as being up-to-date. LEA directories should include some usable information such as enrollment; number of schools by type or grade span; number of teachers; and expenditures per pupil.

Enrollment and Attendance. As mentioned elsewhere, the current NCES data program presents some problems of terminology or nomenclature which should not be entirely overlooked. It is difficult to think of primary and pre-primary pupils as students, which is a term traditionally reserved for the college attendee. Likewise, fall membership leaves something to be desired when referring to enrollment on, or near, October 1. Membership has been a term denoting the average number of enrollees over some period of time, and it has been defined in the handbook series. At any rate, accurate enrollment by grade and level, collected annually (periodically by sex and age) is vital to the statistics program. Average daily membership (ADM), and average daily attendance (ADA), are measures of pupil load which are not now comparable State to State and not available in some areas. These are things the Administrator and the Chief State School Officers should attempt to define and make universal. Leadership is vital and, in some cases, State legislation would be desirable, e.g., ADA in California. It is amazing that California has not corrected this unusual situation of allowing pupils, with valid excuses, to be counted as present. Over the years, it has cost the State school system millions in federal aid money - Elementary and Secondary Education Act, Title I (later called Chapter I), and School Assistance in Federally Affected Areas (SAFA or Impact Aid). ADM and/or ADA are the only measures of pupil load which will make it possible for researchers and administrators to make the needed comparisons.

Class size is another statistic needed periodically to show the distribution of the pupil burden on the individual teacher. Pupil/teacher ratio is not a substitute for class size and the idea of making that substitution should be discouraged. A periodic survey involving a small sample would produce a distribution of class size by State, and need not be done more often than every three to five years. It has not been done adequately for many years.

Several other sets of numbers should be made available on a regular basis through sample surveys or updates: length of the school term, compulsory attendance ages, pupils transported, and enrollments in various programs or subject areas.

It is possible that some of the items mentioned here could be obtained by the Bureau of the Census; for example, enrollment by grade, sex, and age.

Employees

Administrators. Professional administrators should be accounted for in at least two groups -- those in the central office, and those whose activities and locations are school-centered. Full time and full-time equivalent of part time would be needed, as well as the salary expense for the two basic groups. Adequate information should be available from a biennial sample survey which would produce a distribution. The periodicity could be longer if experience shows this population to be stabilized.

Teachers. The number should be collected by a sample survey which would produce a State-representative distribution by employment status -- full time, and full-time equivalent of part time. The associated salary expenditures should be collected biennially to produce good figures on average salaries of teachers. As mentioned elsewhere in this paper, attribution of some salary-related items will be required in a few States. Additional sample surveys at intervals of about four or five years should collect data on teachers by sex, by assignment and level, by training, and by years of teaching experience. It is conceivable that these data could be collected by some other agency such as the National Education Association, but they should be institutional data as are almost all those discussed in this review.

Other Professionals. Data are needed on other professional employees such as guidance workers, psychologists, and librarians at the same level of detail as for teachers.

Non-professional (Classroom associated). Data on the non-professional employee in the classroom should be approximately the same as for the teachers. Using a sample may be risky, so it may be imperative to include the necessary items in some universe survey, or a specially designed sample.

Non-professional (Non-classroom associated). Data are needed periodically on the number of full-time equivalent non-professional employees by general assignment (office-clerical, maintenance, cleaning, bus driver, etc.). Adequate data should be obtainable from a small sample after gross figures have once more been obtained.

High School Seniors, Graduates, and Dropouts. Data on the study programs of high school seniors, at least in general terms, should be collected and analyzed on a periodic basis. The number of graduates and the general area of their studies should be made available by State, and by sex. General Educational Development (GED) certificate recipient data must be reported separately from those who receive regular diplomas.

Properly designed dropout studies should be done at regular (3 - 4 years) intervals through the use of a sample which will estimate a State distribution by size of school system. Reliance on the retention rates, presently the only substitution for dropout rates, indicates a reluctance to attempt a difficult task even when there is a clamor for better data than that produced by NCES.

Revenues and Expenditures

Revenues. Income or revenue by source (Federal, State, intermediate, local, and other) should be collected and reported by State. A distribution by size of school system would be a benefit, even if available only every four years. If State aggregate figures are not easy to obtain on an annual, routine basis, a sample study every other year should be adequate. There should be some exhibit items such as the proportion of local funds from property taxes. Dependent districts and those States where unusual financial arrangements prevail will require that some attribution be done by a knowledgeable educational statistician. For instance, if the State makes contributions directly to the teacher retirement system, those amounts must be attributed to the salary expense item, and to an appropriate revenue item.

Non-revenue receipts should be available by State. This is a necessary item, but in some cases so much attribution is required that the figures should be collected biennially or less frequently unless some of the problems can be solved. If non-revenue data are not obtained, the total picture will not be available.

Expenditures. The various financial accounting handbooks and their revisions have introduced some confusion into an area of school statistics where there was less than total agreement before; however, even the most recent revision allows the collection of needed data since there have been no major changes introduced by the 1980 revision. In a separate section of this review, there appears a set of comments on the NCES Common Core of Data (CCD) which impinges on the following list of items that should be available in all school systems using the new handbook. Some combining of items will be required to make comparisons with States where the new handbook is not being used.

Current Expenditures
  Elementary/Secondary Instruction Programs

  Support Services, Instruction
    Attendance
    Guidance
    Health
    Psychological Services
    Speech Pathology
    Instructional Staff
    Other Instruction

  General Administration
    Business Services
    School Administration
    Operation and Maintenance of Plant
    Transportation

Facilities Acquisition
  Capital Outlay for Equipment, Buildings, and Sites
  Interest on Long-Term Debt Associated with Building
  Repayment of Principal of Long-Term Debt

Projections. Most of the items detailed here should be included in a complete set of projected statistics. The information people must have projected statistics -- some of their "customers" will not take "no" for an answer. Suppose the White House asked for an estimate of expenditures for public schools in 1988-89 and the answer, in due course, went back. Then the same questioner might say he wanted an estimate of elementary/secondary enrollment for the same year. These answers better "fit together" or there will be embarrassment enough for everyone. Without a coordinated set of projections, these questions will probably not be handled correctly.

The information staff can do a great job of estimating statistics for the current year and maybe the next year or two, but beyond that, too much demographic information is needed for them to have a good answer for 1995-96, for instance. Restarting the program would require some time and resources, but would be a worthwhile expenditure.

Outcomes

Outcomes. Measurement of outcomes is a morass which should be avoided; however, many inquirers want to know how a particular school, school district, or State school system ranks with others.

The Federal or State governments probably should not design a testing program to make these outcome comparisons. Those States with graduation test requirements typically agree that individual, school, or school district data will not be made public. How could a national achievement test be given? States have shown interest in expanding the sample for the National Assessment of Educational Progress (NAEP) to obtain State representative figures; however, these are not readily made available to the researchers or news media. They don't mean much unless State estimates are obtained for all States.

It may be feasible to equate existing test results through the techniques used in the Anchor Test Study, done several years ago by Dr. Charles Hammer (NCES), where results of eight reading achievement tests were equated so that one could find a comparable test score for any of the eight tests when compared to any other. This project was successful, but has not been widely used. In fact, it has been used hardly at all. Could the States agree on a set of achievement tests that would be administered by each state, or could all those that exist be equated; it is doubtful.

This is the kind of thing the Secretary could get his teeth into and it might even work for him, but it is recommended that the Center remain ready to help but not lead.

Comments

Questions and Comments about CCD. The word "membership" has always been connected with some sort of average, such as average daily membership. What then, is fall membership as in Part I, Public School Universe? Obviously, it is the number enrolled (on or near October 1). Why not ask the schools for the number of full-time classroom teachers, and the full-time equivalent of part-time classroom teachers? The response burden would not be increased because these numbers are typically known separately. The data tell us something about employment practices, particularly if compared over a period of time. All of the Part I data would be useful in a set of school directories which have not been published for about seventeen years. The data are available from computer files, but that is a poor substitute and is not widely known.

Are all the items needed for sampling available in Part II, LEA Universe, or a combination of Part II and Part III, LEA Non-fiscal Report? In Part III, LEA Non-fiscal Report, there are more problems with nomenclature: does instructional staff include guidance personnel, etc., as has been traditionally true? Student membership is used again when apparently enrollment is what will be given by the respondent. There may be a problem in Part III, LEA Non-fiscal Report unless the question about enrollment is asked more clearly than it is stated in the list provided. The attempt here seems to be to get at part-time attendance of children below the first grade. It's better to ask how many there are, and the length of their school day, then do your own arithmetic, than to wonder if it has been done, or done correctly. A regular program for publishing a directory of the Part II data every third year would be a good service. It would be a better service if a little more information could be added, such as fiscal status and current expenditures per pupil.

Part IV, LEA Finance Report, is the worst example in CCD; it shows a lack of understanding of the problems and presents unworkable solutions. It is not necessary (or desirable) to obtain these items every year. A relatively small sample will produce a distribution suitable for most needs, most years, with a census tabulation periodically (three to five years). The construction of the data set seems to assume that the State and local education agencies are using the Financial Accounting for Local and State School Systems manual and that there are severe limitations on the data items available. First, relatively few State and local systems are using the manual. Second, there is no reason to omit the many functional accounts which have been used since 1957, except that the old Auxiliary Services, or Fixed Charges account no longer exists -- the amounts formerly collected there are now distributed to those accounts with which they are associated. Functional account data should be available since Mr. Barr says in the Financial Accounting foreword, "...it does not make major changes in the account classification system." The survey director will need to know (as will users) what is included in such classifications as Instruction, Attendance and Health, Operation and Maintenance of Plant. Putting these accounts together is not very new and causes researchers no serious problem.

Under other uses of public funds, more items are needed. For instance, it is not enough to ask for debt service, it is necessary to know interest on long-term debt and repayment of principal separately. If construction equals capital outlay, there is no problem; but does everyone understand this?

There seems to be enough similarity in the State aggregate and LEA items to make a charge of duplication, yet there are needed items in both. For example, State aggregate current expenditures for, or on behalf of, LEA's should be attributed to the proper program in the LEA, so that the resulting figures will be like those in other districts. Examples of these include State contributions to teacher retirement funds, State expenditures for buildings, and financing of debt. Comparable attributions to revenue accounts must also be made.

Comments on Sample Surveys. The sample surveys present a pretty picture, but so far the private school survey has not produced usable data. It appears to be too ambitious when the small proportion of children enrolled in private schools versus those in public schools is considered. To a person with little knowledge of sampling, the sampling fractions for public and private schools seem disproportionate. It is difficult to see what will be learned from the teacher demand and shortage survey unless there has been much work on definitions. For example: respondents should not consider a vacancy filled if a make-shift arrangement has put a poorly-qualified and poorly-motivated teacher in a classroom where a better trained individual is needed. Some of these surveys have produced valuable data, but an estimate of the distribution by State is needed in many instances.

Comments on paper by Cooke, Ginsburg, and Smith. The tone of "The Sorry State of Education Statistics" strongly implies that the ills of education statistics are the fault of things done by NCES along with some things not done. Unfortunately, there is some truth in what they say; however, many of the specifics are wide of the mark. Example: NCES does not have dropout statistics, but there is no suggestion that the retention statistics represent a substitute. Bureau of the Census data are presented in the Digest. Example: NCES has not reported on class size for many years; the data shown in the Digest are pupil-teacher ratios which are not claimed to substitute for class-size data. Example: Advocate groups such as those who favor bilingual education tend to overstate the size of their group which may well be enough to explain the differences that Cooke, Ginsburg & Smith complain about. Example: When one says that there is a 300 percent difference between 1.2 and 3.6, anything else said becomes suspect.

In Conclusion

This review does not present a detailed list of data items, nor does it recommend a data gathering plan, but it presents some ideas and recollections of problems and needs in providing information service over a number of years. Even to begin to do what has been suggested here, NCES would require support from Congress, from other users, and from the States. Additional and/or different staff would be required to do some of the work -- particularly the research and attribution suggested here.

Research on the measurement of outcomes should be increased; perhaps the higher education community could be helpful, or perhaps the Chief State School Officers or others may have expertise to lend to this huge, difficult, sensitive task.

It is difficult, if not impossible, to guess what the issues and data needs might be in 1999. If the trend toward greater interest in schools -- particularly by parents -- continues or increases, many factors discussed here could change drastically. Not the least of these would be interest in studies of class size and average salaries of teachers, as well as items which describe teacher fitness -- years of teaching experience, training, test scores, etc. Parents are not generally impressed by top-heavy administrative staffing, or inflating school bureaucracies, or the accompanying salaries. Even if parents don't become more interested in schools, the data discussed in this review will continue to be needed and should be provided by the only appropriate agency, the National Center for Education Statistics.

EDUCATIONAL STATISTICS AND SCHOOL IMPROVEMENT

by

Willis D. Hawley

STATISTICS AND THE FEDERAL ROLE IN EDUCATION

The collection and dissemination of educational statistics have been a federal responsibility for almost 120 years. The federal role in education has changed considerably since then and the demands made for the scope, depth and accuracy of information have increased accordingly. While the appropriateness of the federal government's role in the collection and provision of statistical information is not seriously questioned, debate about what and how data should be collected is recurrent. The reason for this is that statistics are not ends in themselves—they support and facilitate the attainment of other purposes. Many of these purposes are controversial and involve real or potential conflict not only within the federal government but among different levels of government. There seem to be four broad purposes for which educational statistics are used:

1) To provide information on conditions and trends relating to the characteristics and performance of learners and educational institutions,
2) To guide the allocation of resources,
3) To assist in the enforcement of laws and regulations assuring students freedom from racial, ethnic, gender or other invidious discrimination, and
4) To facilitate the improvement of educational policies and practices.

Until recently, most uses of educational statistics related to the first three of these purposes. However, the current surge of interest in educational reform has seen policymakers and reformers turn to educational statistics to determine how effective schools have been and what improvements might result in the greatest increases in student learning. And, as more and more people have sought to put educational statistics to work in the cause of change, the limits of available information have become more apparent. As Cooke, Ginsburg and Smith (1985) point out, information available from the National Center for Educational Statistics (NCES) is often inaccurate, incomplete and inconsistent. NCES has no monopoly on such credits. In particular, the extant information tells us too little about the outcomes of education. And, when such information is available, measures that might account for differences among students and school systems often are not.

This paper focuses on how educational statistics might better serve the quest for educational improvement in elementary and secondary schools. In seeking answers to this question, I briefly identify the contributors to the federal information base and suggest that available resources could be better used. I then turn to the notion that if we want statistics to serve the interests of educational improvement, we need to conceptualize the sources and processes of school productivity. Such an analytical model is presented and the types of questions suggested by this framework for which statistical information could provide answers are identified.

Some implications of this approach for current studies undertaken by federal agencies, especially the National Center for Educational Statistics (NCES), and for future inquiries, are explored. I conclude by arguing that the quality of educational statistics can be improved if (a) statistical surveys are theory driven and policy focused and (b) the use of statistics is made easier and more productive.

THE RELATIONSHIP BETWEEN EDUCATIONAL STATISTICS AND EDUCATIONAL RESEARCH

The collection of educational statistics should both be informed by and facilitate research. Statistics, by themselves, seldom tell a very rich story but they can and should encourage better story telling.

I use the term research here to mean the process by which new knowledge that explains why things happen is discovered. To believe that one can understand what is going on in schools by knowing the characteristics of schools, even if one knows the performance of students in those schools, is to be presumptuous or naive--or both.

Statistical gathering and analysis that is uninformed by research encourages simplistic conclusions about the causes of student performance. There are two reasons for this. The first is that important variables may be omitted from the data sets while irrelevant data may be collected. Requests for information that few people will or could use makes unproductive demands on data providers but the omission of variables encourages inappropriate analysis. A second characteristic of statistical information the collection of which is not informed by research is that key variables may be inappropriately specified and measured.

Data that have been shaped by research and collected with theoretical concerns in mind obviously facilitate research. Both the extensive utilization of the National Longitudinal Study of the High School Class of 1972 (NLS-72) and the limitations of these analyses makes this point. In comparison to the NLS-72, the High School and Beyond (HSB) Study was much better informed by research and theory and the potential payoffs from analyses of these data appear to be enormous.

Financial resources available for educational research are meager. At the same time, there is a virtual army of actual and potential educational researchers that could attack "research-influenced" information on education in search of new knowledge. There are ways that the energies of this army could be enlisted and redirected from the ubiquitous mail surveys and meaningless descriptive studies but an exploration of such strategies seems to be the topic of another paper.

Research based on the types of statistical information typically collected by government agencies or their contractees, even data as rich as those collected in the HSB Study, can tell only part of the story. Analysis of large scale statistical data should lead to and be informed by more intensive research that examines the dynamics of student and teacher interaction and otherwise helps us to avoid false assumptions about the meaning of statistical data. For example, the introduction of comprehensive teacher evaluation plans, especially those that involve standardized statewide criteria and processes, is one of the most far-reaching reforms ever introduced in American schools. But it is not the presence of these plans or even the putative characteristics of the plans that will explain any changes in teacher behavior; it is the way these plans are implemented and the data from them are utilized that will tell the tale. One cannot determine how an evaluation system (a curriculum, or other innovations such as a new instructional method, or a management practice) is implemented and how context affects implementation without using research techniques that are more intensive than statistical surveys.

In summary, research and statistical surveys should be seen as having a symbiotic relationship. This is particularly true with respect to statistics that might inform decisions about school improvement.

THE UNDERUTILIZATION OF EXISTING CAPACITY

The National Center for Educational Statistics collects, directly and indirectly, only a portion of the information that might be relevant to educational improvement. Within the Department of Education, a host of other data collection efforts are regularly undertaken. These include various policy studies and program evaluations conducted or supported by the Office of Planning Budget and Evaluation; research and information collected by line agencies such as the Office of Special Education and the Office of Bilingual Education and Minority Language Affairs; the National Center for Research on Vocational Education and the National Institute for Research on the Handicapped; The Office for Civil Rights; and various programs within the National Institute of Education, including the National Assessment of Educational Progress (NAEP). This list is by no means exhaustive. The point is that an enormous amount of information is collected on American education but there is no central effort to plan or coordinate the information collected or even to consolidate it once it is collected. This is not to argue that all data collection efforts within the Department of Education should march to the same drummer. But, the virtual absence of past efforts to define the information needs of the educational improvement mission of the Department of Education denies the Department opportunities to inform the country and influence policy. Various mechanisms have been employed over the years to "control" data gathering efforts but these have been focused primarily on the reduction of paperwork and other burdens imposed on data providers rather than on improving the quality and accessibility of information.

Given the absence of efforts to plan and coordinate data collection and to consolidate statistical data within the department, it is not surprising that the education-related data collection efforts undertaken by other federal agencies are not part of a comprehensive plan and that data are not assembled in one place. Among the agencies that regularly collect information that might be helpful in understanding and facilitating school improvement are the National Science Foundation, the Department of Labor, the Department of Health and Human Services, the Department of Commerce (especially the Bureau of the Census), and the Department of Defense.

The Secretary of Education could take the initiative in designing a master plan for education-related statistics that would encompass the statistical activities of all of the agencies identified above. The first step in that regard would be to catalog current and planned programs. The second step would be to identify the key variables upon which major studies focus and the uses to which the data are part. A logical outcome of such an initiative would be the establishment of a Federal Interagency Advisory Committee on Educational Statistics. Such a committee could be staffed from the Office of Educational Research and Improvement. An interagency effort could (a) identify sources of data (b) suggest how existing data can be integrated (c) identify areas of unnecessary hindrance and important issues about which data are needed and (d) provide advice to the developers of major new efforts to collect educational information. * One difficulty that will confront such an effort is that the Department of Education is a minor player in the statistics collection game; it does not have a lot to trade with when it seeks cooperation. But the interest in education is great now and, at least for the short term future, the possibility for collaboration may exist.

---

* NCES has regularly included representatives from the Census Bureau and the Bureau of Labor Statistics on its advisory board and this seems to have been constructive.

States and localities generate much of the data collected by the federal government. In addition, states and localities collect a great deal of information that does not now become part of the federal data base. The growth in the capacity of non-federal education agencies to generate and analyze data appears to have been extensive in the last few years. These "unassembled" data have potentially significant uses in fostering school improvement but this potential is undermined by the absence of standardized definitions of key variables, variation in the information collected from state to state, differences in data collection processes and the difficulty of retrieving the data collected.

A good example of the weaknesses of current state level data is provided by information about teacher supply and demand. At a time when teacher shortages are widely anticipated, states presumably would find it valuable to know how their policies shaping supply and demand compare in effectiveness to those of other states. Some comparative data on state policies are available (cf. Roth and Mastain, 1984). However, data on the consequences of these policies—e.g. number of certified teachers seeking employment, the qualities of these teacher candidates, who actually enters the profession, the volume of reentry, and the rate and character of teacher attrition--are hard to come by even though such data are available for states to secure.

The federal government could take the lead in encouraging and even assisting in the collection of policy-relevant information on the supply and demand for teachers. One mechanism for planning such an initiative is the not overworked Intergovernmental Advisory Council on Education. NCES could be responsible for implementation.

A THEORETICAL FRAMEWORK FOR SPECIFYING STATISTICAL DATA RELATED TO THE IMPROVEMENT OF SCHOOLS

Like data analysis, data collection related to school improvement should be guided by two central questions:

1. What are the problems or issues we want to understand better?
2. What theory or theories might best identify the range of factors that influence the outcomes in which we are interested?

If the answer to the first of these general questions is to know how to most cost-effectively improve schools, the answer to the second question depends on how we would decide a school had improved. I suggest that the central purpose of schools is to produce student learning. Thus, improvement would be measured in terms of amounts or rates of learning or, at least, in terms of the relative achievement of students once all of the variables schools cannot influence are taken into account. I will return below to the types of learning in which we might be interested.

OVERVIEW OF THE MODEL

The effectiveness of most organizations that are in the business of producing something is the consequence of the interaction of five types of factors:

- the clarity and specificity of goals being pursued (e.g. student academic achievement),
- the nature of the raw materials (e.g. the potential for learning students bring to school),
- the sophistication and appropriateness of the technology or means of production (e.g. the curriculum), and
- the quality of the craftmanship (e.g. teacher behavior)
- "environmental" conditions that facilitate or impede productivity

Within this framework, the specific variables to be measured will have to be identified. I will suggest a number of the key factors that current research suggests accounts for variation in student achievement from classroom to classroom and from school to school. This model is based on an extensive recent review of research syntheses and individual studies (Hawley and Rosenholtz, 1984). The relative importance of these factors for enhancing student achievement has been affirmed by a report urging school reform prepared by thirty-two school superintendents from around the country (National Consortium on Educational Excellence, 1984).

There are, of course, other efforts to synthesize current research on school effectiveness (cf. MacKenzie, 1983; Brophy and Good, forthcoming; Purkey and Smith, 1983; Northwest Regional Education Laboratory, n.d.). While there is substantial overlap in the conclusions reached in these reviews, there are also differences. Moreover, the model presented below is derived from research on student achievement measured in conventional terms (i.e., tests of skills and knowledge) and that is a significant limitation. The "Learning Productivity Model" outlined here should be considered illustrative. It seems to be, however, a useful starting point in identifying the types of statistics that will be most helpful in understanding and facilitating school improvement.

Because scholars disagree about the major determinants of school effectiveness and because the importance of different factors depends on the outcome of schooling with which one is concerned, the model or models which would guide federal data collection and coordination efforts should be derived from a consensus of leading researchers and practitioners. The process by which such consensus might be developed is not difficult to imagine: experts should be selected by a snowball reputational technique and they should be provided with alternative assumptions about variables and their interrelationships. The process should be interactive until agreement is reached either about particular variables or alternative explanations for specified outcomes. Thus, I have chosen not to try to detail specific variables about which one might want statistical data but instead to pose questions which direct attention at the types of data that would be useful in assessing and fostering school improvement.

The Learning Productivity Model of school effectiveness focuses attention on schools and classrooms. This seems an appropriate strategy for organizing one's thinking about school improvement given the substantial evidence that change that affects the quality of education children experience must be generated at the school level if it is to have a significant and continuing influence (Hawley and Rosenholtz, 1984, Ch. 1). Moreover, the frequency with which one sees schools of widely different quality within the same district, even if one takes into account differences in student body characteristics, reinforces the idea that school improvement-related data should provide well textured pictures of schools. However, a number of environmental conditions create the context for school improvement and need to be considered.

The types of factors that influence school productivity and their dominant interrelationships are summarized in Figure 1. Table 1 identifies key variables that comprise each of these factors. Before turning to a discussion of these variables and the directions they point to with respect to the collection and provision of statistical information, the issue of educational outcomes needs to be addressed.

The national interest in school improvement obviously requires that we focus on the ability of schools to produce certain outcomes in the form of student learning. I want to distinguish between these school outcomes from the consequences of formal learning. I will refer to the latter, for lack of a better term as "lifetime outcomes".

SCHOOL OUTCOMES. There appears to be widespread agreement on the importance we should attach to the acquisition of so-called "basic skills" and knowledge about the more or less traditional subject matter measured by tests like the National Assessment of Educational Progress, the Achievement Tests and Advanced Placement Tests administered by the Educational Testing Service, and various standardized tests which report student performance in terms of grade level equivalencies.

There is growing concern, however, reinforced by the poor showing of students on tests of so-called "higher order skills", that our schools are failing to teach students to reason inferentially, to engage in reflection about and systematic analysis of complex problems, and to be creative in the face of uncertainty. Assumptions about the importance of these and similar capabilities to the life chances of individuals and to the economic and social health of the society have yet to be tested. But, there can be little doubt that interest in these types of school outcomes is growing. There can also be little doubt that there is an absence of consensus not only about how to measure these outcomes but how to talk about them with shared meaning. It follows that little is known about how variations in curricula and instructional practices relate to these outcomes. Thus, the federal government could contribute to enlarging the definition of school effectiveness if it could facilitate the development of clearer conceptions of the outcomes by which the effectiveness of schools might be measured. The importance of such an effort, which could take the form of conferences and relatively simple studies of expert and leaders' opinions, is significant because it is possible—as Cooke, Ginsburg and Smith (1985) observe—that the factors that maximize student performance on tests of basic skills may not produce other types of student learning, such as "higher order skills". Indeed, some researchers believe that some elements of the technology of schooling that produces learning of the so-called basics may constrain other learning.

LIFETIME OUTCOMES. Educators, policy-makers and parents have debated the priorities that schools should emphasize since there have been public schools. The nation seems to swing between different priorities with the various advocates invariably making assertions about the effects that one focus or another will have on the experiences and well-being of students once they leave school (cf. Ravitch, 1983). Almost all of this advocacy, however, is unburdened by evidence about the relationship between variations in school experiences and differences in students' post-school outcomes.

While it might be argued that curricula should reflect the values of society, or the community, or the parents independent of the extrinsic consequences they might have for students, this is an unsatisfying position seldom asserted by the advocate of different curricular emphases. Thus, it seems appropriate that the information we have about America's educational systems include data that would allow us to understand how differences in content and levels of investment lead to differences in outcomes beyond academic performance and educational attainment.

Two obvious educational outcomes that it seems useful to know about are occupational success and income. The importance of such information is underlined by a recent synthesis of studies that concludes that individuals' academic performance (as measured by grade point averages or standardized test scores) have almost no relationship to occupational success (Samson, et al., 1984). The central theme of this review of research is reinforced by another recent study showing that a large national sample of employment officers place like emphasis when making hiring decisions place little emphasis on most of the academic outcomes of education that schools are now being asked to focus upon more intently (Crain, 1984).

Other nonacademic outcomes of schooling that might help resolve some of the recurrent debates about curricula and thus focus reforms on courses of study which benefit students are:

- participation in the political and social life of communities
- incidence of antisocial behavior
- family stability
- condition of physical and mental health

The importance of considering the goals of schools and the priorities attached to them is critical in understanding the potential for school improvement in any given school or school system. The current educational reform movements, for example, rests on the assumption that the quality of schools has declined. The evidence supporting this assumption is conflicting, however, (cf. Smith, 1984, Hawley, forthcoming) if we focused attention on how well the schools did between 1970 and 1980 in (a) promoting basic skills, (b) meeting the special needs of the disadvantaged and the handicapped, (c) increasing the number of students who did well on Advanced Placement Tests, and (d) reducing the dropout rate and increasing the number of students who attend at least some college, the schools would have to be judged moderately effective. If we ask how well schools did in teaching inferential reasoning and science, or developing students' understanding of democratic process, most schools would have to be judged to have declined in effectiveness. If the explanation for improvements and declines have to do with how schools allocated time and other resources rather than how well they used them, the implications for school improvement would obviously be very different. This rather simple notion that the current dissatisfaction with schools is the result of increased and/or different expectations rather than declines in school quality has received almost no attention in the contemporary public discourse about education in part, perhaps, because there has been little evidence on how school outcomes are related to differences in the priorities given to particular goals and the way these goals have been pursued.

Recognizing that the factors that affect school productivity may vary depending on the outcome being focused upon, let me identify the general types of influences that current research suggests determine student learning. The statistical information that seems to be both useful and feasible to obtain about each of these influences is indicated by a set of questions that need better answers than we now have. Recall that the Learning Productivity Model has five sets of interactive variables—goals, raw materials, technology, craftsmanship and environmental conditions. Each of these will be considered in turn. I will not cite the literature that links these variables to student learning, unless otherwise noted. The reasons for believing that these factors and others related to the questions posed are important determinants of school outcomes are documented elsewhere (Hawley and Rosenholtz, 1984).

THE GOALS OF SCHOOL IMPROVEMENT

RELATIONSHIPS BETWEEN GOALS AND OUTCOMES. Organizational goals typically focus on products or processes. For our purposes, product goals—the types of learning we want to produce—should be our concern. It might well be interesting to know the fit between process goals and the processes actually being utilized but if we know the latter, which we would if we collected data on the elements of The Learning Productivity Model, we will have this type of statistical information about the processes that are relevant to school improvement.

THE LEARNING PRODUCTIVITY MODEL - KEY FACTORS AFFECTING SCHOOL EFFECTIVENESS

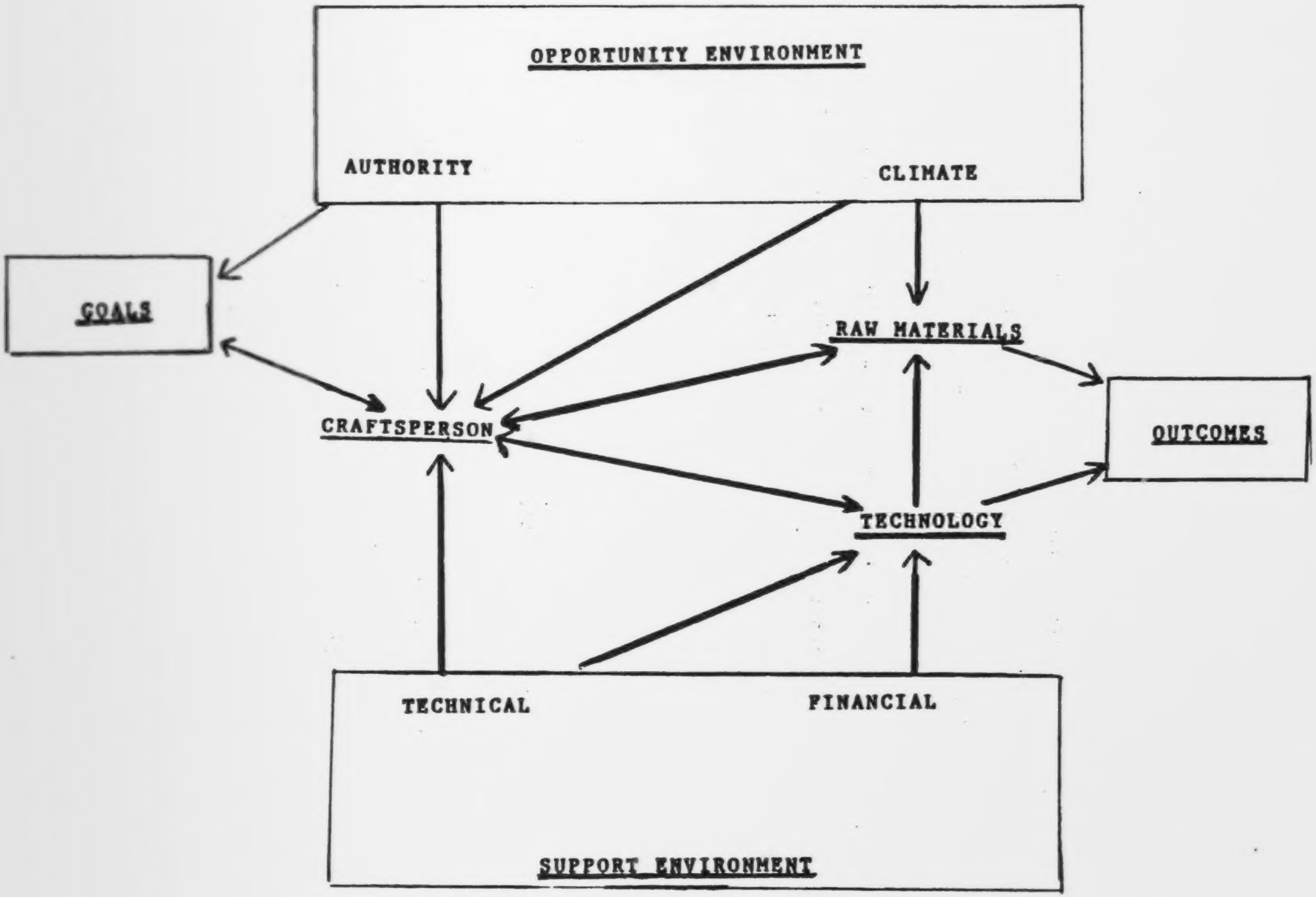

OPPORTUNITY ENVIRONMENT
AUTHORITY
CLIMATE
GOALS
RAW MATERIALS
CRAFTSPERSON
OUTCOMES
TECHNOLOGY
TECHNICAL
FINANCIAL
SUPPORT ENVIRONMENT

TABLE I

KEY ELEMENTS OF THE LEARNING PRODUCTIVITY MODEL OF SCHOOLING

GOALS

1. CONTENT
2. CLARITY
3. FOCUS
4. CONCENSUS

CHARACTERISTICS AND QUALITY OF RAW MATERIALS/STUDENTS

1. COGNITIVE CAPABILITIES
2. ACQUIRED KNOWLEDGE
3. MOTIVATION
4. OUT-OF-SCHOOL RESOURCES FOR LEARNING
5. OUT-OF-SCHOOL CONSTRAINTS ON LEARNING

THE TECHNOLOGY FOR PRODUCING LEARNING

1. INSTRUCTIONAL STRATEGIES
2. ORGANIZATIONAL ARRANGEMENTS (EXAMPLES: CLASS SIZE,
3. CURRICULUM
4. TIME
5. LEARNING RESOURCES (EXAMPLES: TEXTS, MATERIALS, TECHNOLOGY, ETC.)

CRAFTSMANSHIP

1. ABILITY AND COMPETENCE
2. CONDITIONS THAT FACILITATE THE USE OF COMPETENCE
3. MOTIVATION

ENVIRONMENTAL CONDITIONS

1. OPPORTUNITIES
2. RESOURCES

One rather straightforward definition of an effective school is that it is a school which achieves its goals. In this sense outcomes and goals can be the same. In order to understand the sources of school effectiveness, however, it will be useful to know the characteristics of the goals to which school personnel subscribe.

CHARACTERISTICS OF GOALS. There are at least four characteristics of the goals to which individuals in a school might profess commitments that appear to be theoretically related to productivity. The first of these is the intensity with which people subscribe to particular priorities. For example, two people may say that reading achievement is their first priority but one may feel that reading supersedes all else by several orders of magnitude. For another, both reading achievement and math achievement may have high priority even though math is ranked lower than reading. Saying that both individuals give their first priority to reading is misleading. Most ways of assessing priorities, however, use simple ranking procedures. (What is needed here is something like the "temperature scales" developed in the voting studies conducted by the Institute for Social Research at the University of Michigan.)

A second dimension of goals that is relevant to student learning is the clarity with which their advocates can describe them. How precise, in other words, are the outcomes being sought? Do seemingly similar goals subscribed to by different teachers have the same meaning for each teacher when the specifics are elaborated?

A third aspect of goals about which data might be gathered is focus. Theodore Sizer may overstate it when he says that the three most important things in explaining school effectiveness are focus, focus and focus. But the point is well taken (cf, Peters and Waterman, 1983, vis-a-vis the importance of focus in private organizations.) In other words, how many high priority goals do teachers and administrators in a school want to accomplish?

A fourth characteristic of goals which appears to influence school effectiveness is how widely shared the goals are.

THE RAW MATERIAL OF EDUCATION - STUDENTS

Students are the raw material with which schools work. Students bring with them to school different capacities, knowledge and motivation for learning and the effectiveness of schools cannot be measured unless these things are taken into account. Moreover, schools can shape their probabilities of success by influencing the learning readiness of students. Ideally, we would want to know something about the intellectual capacity of students but this is difficult to assess validly under any circumstances and seems beyond the reach of statistical surveys. Other factors that schools might influence or that could be influenced directly by public policies are students' previous knowledge, students' motivation, the resources and assistance students have available out of school, and home or neighborhood constraints on learning. Thus, we might better understand the effectiveness of schools and the promise of different improvement strategies if we had data that spoke to the following questions.

What was the performance of the students in their previous school or class with respect to the outcomes in which we are interested? What is the socioeconomic status of the students? What types of preschool learning experiences, if any, did students have? What types of learning-relevant interactions do students have with their parents, older siblings and other adults with whom they spend considerable time? What proportion of the students are from single parent families and how many have a parent at home when school is out? What does the school and teachers do to involve parents in the education of their own children? How much and what kind of homework do students do and under what conditions can they study at home? What kind of supports for and constraints on learning are there in the students' immediate neighborhood? How much time do students spend watching television and what do they watch?

THE TECHNOLOGY OF EDUCATION

When students go to school they experience a range of structures, processes and learning resources. These "means of production" --which obviously vary considerably among states, school districts, schools and classrooms--make up the technology of schools. Of course, the effectiveness of the technology in producing student learning is significantly influenced by how well it is used but we can make policy—and do—about what I am calling technology more or less independently of the policies we make that affect the quality and performance of teachers and administrators.

Data collected by NCES and other federal agencies are typically short on information about what actually happens in schools. The HSB data go a long way in addressing this limitation of available statistical evidence but, of course, they tell us only about high schools at two points in time. The types of information about school technology that research suggests would be helpful in assessing school productivity and developing improvement strategies include data on instructional strategies, organization arrangements, curriculum, the use of time, and learning resources. Some of the questions we would want this information to address are set out below.

INSTRUCTIONAL STRATEGIES. What mix of strategies--such as whole class teaching, programmed instruction, individualized learning, cooperative learning, mastery learning, peer tutoring and interactive teaching--is employed and how does this vary by subject matter and learning objectives? How often are students evaluated and how are decisions made about movement through the curriculum and grade levels? How high and how clear are standards of student performance and what types of rewards do students receive for meeting standards?

ORGANIZATIONAL ARRANGEMENTS. What is the class size with and without teacher aides? How many students are there in the school and how are they assigned to instructional units (e.g. grades or "houses" or blocks)? What grouping and tracking procedures are employed? What is "skewness" of student performance in particular classrooms?

CURRICULUM. What subjects are taught, at what level of difficulty and for how many hours each day or week. Does the school have a core curriculum that is well articulated across grade levels? Are the tests used to measure student performance articulated with the curriculum?

TIME. How many hours of actual instruction are allocated each week free from interruption? How many hours do teachers teach each week? How many days each year do students attend school?

LEARNING RESOURCES. What is the condition and nature of the instructional facilities? What support systems do teachers have—library, volunteers, audio-visual, business-school linkages, etc.—and how often are they available and used? How many computers are available and how are they used?

THE QUALITY OF THE CRAFTSPERSON

A decade or so ago, researchers and policymakers--but not parents--asked questions like "do teachers make a difference"? The recent outpouring of research on school effectiveness leaves little doubt that both teachers and administrators make a big difference in the quality of education students receive. (cf. Hawley and Rosenholtz, 1984, Chs. 1-3). For example, not only do teachers influence how students learn by the way they implement a curriculum, teachers also shape what students are taught (Brophy, 1980; Green and Harker, 1982). Recent research also suggests that teachers significantly influence parental involvement in their children's education (Epstein, 1984).

Statistical data about educators is very limited. It is not possible, for example, to get a clear idea of what the career paths of principals and administrators look like under different circumstances. Little is known about the qualities of those who actually teach or what difference these qualities make in student performance (cf. Evertson, Hawley and Zlotnik, 1985). Very little is known about teacher attrition. And so on. (The current activity within NCES to develop a better picture of teachers is ambitious and well conceived and it will be much welcomed by policymakers. A similar effort with respect to school administrators is also needed.)

There are three interrelated aspects of craftsmanship which appear to influence student learning: ability and competence, conditions that facilitate or impede the use of competence, and motivation. Some questions to which it seems important to have answers about these interrelated aspects craftsmanship follow.

ABILITY AND COMPETENCE. What are the levels of general academic ability and subject matter knowledge that teachers and principals possess? How much of what types of formal training have teachers and principals had and at what stage of their career did this training occur? How many and what kind of opportunities for increasing their professional competence have teachers and principals had? What opportunities have teachers and principals had to learn informally from their peers?

MOTIVATION. What priorities do teachers say they give to the different reasons why they teach? What is the volume and character of feedback about their performances that teachers and administrators receive? Is there a formal evaluation system? What are its criteria and how is it used? What is the salary range that is accessible and on what bases are salary increases awarded? Are there financial incentives available other than salary increases? Do teachers and administrators have an opportunity to participate in key decision affecting professional practice? In what ways is superior professional performance recognized and rewarded (besides economic rewards)?

CONDITIONS THAT FACILITATE EFFECTIVE PRACTICE. What level of support do teachers and administrators receive from their respective administrative superiors? How often do teachers have the opportunity to interact professionally with their peers and do norms of collegiality exist? What is the level of order and discipline in the school? Is the teaching time of teachers protected from interruptions and diversion? How widely shared are goals for student learning and at what levels of expectation are these goals pitched? What are the number of students in the school?

It should be noted that NCES has focused increasing attention on teachers in recent months and several surveys have sought to collect data about teachers and what they do. While this effort addresses several of the questions raised above, the bits and pieces apparently cannot be aggregated and too little attention is focused on outcomes. The HSB data do provide student outcome data but little information about teachers' careers and personal characteristics. The NCES labor market survey provides information about teacher shortages and teacher incentive plans but no information about salaries.

ENVIRONMENTAL CONDITIONS

The problem of controlling for non-school factors that affect the productivity of a school cannot be dealt with adequately except under experimental conditions. Therefore, the theory that guides data collection must encompass critical student characteristics beyond the reach of the school and aspects of the schools' ecology. Much has been written about the factors external to individual schools that influence school improvement. Unfortunately, aside from a handful of case studies, little emperical research on the relative importance of these factors exists. Thus, unlike other dimensions of the Learning Productivity Model, this one is not researched-based.

The collection of data relating to school external environments is problematic not only because so little evidence is available upon which to select from among the theoretically interesting ecological factors that could influence student learning but because of the difficulty of acquiring the information. If we were to constrain the collection of data to those that could be provided by occasional national samples or by school systems from existing information bases, the types of information relevant to school improvement that might be collected seems a bit easier to think about.

It seems reasonable to think of environmental conditions that influence school productivity as being of two broad types: (a) conditions that grant or constrain the opportunities of educators and students and (b) resources available with which educators and students can pursue their goals.

THE OPPORTUNITY ENVIRONMENT. The opportunity environment essentially shapes ambitions and expectations and grants—formally or informally—authority to act. Its' dimensions and their consequences might be understood if questions such as the following could be answered. How much decision-making discretion do school administrators and teachers have to make and implement policy? How much support do the goals of school-level personnel have among district-level administrators and other teachers? How stable are residential patterns and pupil assignments to schools? How available to students are low-cost options to pursue postsecondary education? What types of of employment opportunities in the community can students expect to have upon graduation?

THE RESOURCE ENVIRONMENT. The resources available to pursue the ambitions and expectations held by individuals and by the collective professional and student populations of school should affect student learning. Resources, of course, create opportunities but the "opportunity environment" just noted relates to sources of motivation provided by the environment while the "resource environment" provides capabilities that facilitate action relevant to the individual and organizational purposes that have been discussed as deriving from factors that characterize the school and its student body.

The nature of the resource environment might be understood if answers were available to several questions. What is the level of expenditures per pupil taking into account the number of students with needs for resource-intensive programs? What is the quality and quantity of technical assistance available in the district to implement school-level goals? How many individuals participate in school activities as volunteers in support of instructional or extracurricular activities? What resources, if any, are provided to schools by business and public and nonprofit agencies? What is the "tax-effort" (the extent to which the available tax base is utilized) of the state and the school district? This last indicator may be a surrogate measure of the school environment's commitment to public education.

SUMMARY

I have argued that the collection of statistical data related to school improvement should be guided by grounded theories and hypotheses that might explain differences in student learning among schools and classrooms. A conceptual framework implicitly embodying such explanations, which I have called the Learning Productivity Model, has been outlined. Identifying the key elements of models and questions that might profitably guide data collection and analyses does not, of course, add up to specific recommendations about priorities that should be placed on gaining particular information that would facilitate the development of effective school improvement strategies. And, obviously, the ways such variables might be measured are only hinted at in the discussion above. My presumption, however, is that these next steps in the development of a plan for collecting statistical information is given direction by this model building exercise.

CONCLUSION - NEXT STEPS

Let me conclude by briefly discussing two courses of action which would result in statistical data bases that would support the formulation of productive school improvement policies and practices: (a) the development of a plan for collecting new types of data (b) the integration and enrichment of existing statistical information.

THE COLLECTION OF NEW DATA

In general, statistical data related to education that is now available is not particularly helpful in developing new policy. The current array of educational statistics focuses primarily on the condition of education rather than on explanations for that condition. In other words, little effort has been made to relate educational processes to educational outcomes. Moreover, as noted above, when outcomes are assessed, the range of educational outcomes measured is narrow. For example, the 1985 Teacher Questionnaire, while it deals with how teachers spend their time, barely touches on the teacher characteristics, behavior, and circumstances most powerfully related to student learning and provides no outcome data whatsoever about either students or teachers. Of what value is such information? If it does lead somewhere, the probability that it will lead us down primrose paths is at least as great as the prospect that the road it will direct us to will be paved with yellow brick.

This generalization about the atheoretical character of data collection does not apply to all surveys. The most notable exception is the HSB Study. There are lessons that this study has for new data collection efforts relating to school improvement even though that is not the primary purpose of the study. First, schools should be seen as social systems and both formal and informal processes should be assessed. Second, an expert review panel that guides the design and redesign of statistical studies seems to be cost-effective. Thus, the next step would be to convene such a panel, refine the theoretical framework, specify variables, consider the sampling issues and otherwise develop the scope of the project so that the potential benefits and the costs can be estimated. An HSB-type study will be very expensive but the return on investment, in terms of usable knowledge and educational improvement will probably be high, if early evidence on the use of the HSB Study is any indication. Comprehensive data collection efforts like the HSB Study or a similar one which would follow from the Learning Productivity Model would be much more valuable to policy makers and would-be reformers if data were collected that would allow analysts to understand the economic costs of alternative improvement strategies.

Because the statistical error that threatens all data analyses is reduced to the extent that one can take into account all of the alternative explanations for a given phenomenon, and because we typically want large data sets to serve multiple purposes, an obvious problem faced by those who design statistical studies is knowing what data NOT to collect. The potential cost/benefit ratio of data collection itself is one way to make such decisions and theory as well as long run utilization studies can be used to make such calculations. The probable reliability of the data is another factor that should be considered and it would seem useful to develop, perhaps through expert interviews, an understanding of the sources of error in survey responses that could be used across studies. For example, such considerations might include the probability that individuals would perceive themselves or a goal they value being advanced (or the converse) by accurate provision of data and the degree to which the information is readily available.

THE INTEGRATION AND ENHANCEMENT OF EXISTING DATA

I have argued that most statistical surveys provide only part of the picture we need to have of how schools operate and the outcomes they produce. This proposition holds for statistics that address issues other than school improvement. There are good reasons for this having to do with cost, burdens on providers, etc., but among the several studies conducted or sponsored by the federal government, there are many pieces of the puzzle. Moreover, several studies could be enriched if data available from published sources other than statistical studies were added to various data sets. NCES, therefore, might consider the following strategies for making available statistics and recurrently conducted studies more accessible and more usable:

(a) coordinate the data gathering of agencies within the Department of Education, across the federal government, and between federal agencies and other levels of government.

(b) "nest" future studies in such a way that data from the same sites could be integrated. For example, it might be possible to conduct the HSB and National Assessment studies in the same or overlapping locations.

(c) identify common sites at which data has already been collected. For example, the teacher supply and demand study may have been conducted in locations for which data on teacher salaries were collected.

(d) enrich existing statistical data by adding information from other sources. For example, the teacher demand and shortage study could be enhanced by adding information on career entry requirements (available from The National Association of State Directors of Teacher Educational Certification).

Steps such as these, coupled with efforts to prepare the data in formats that are well documented and readily usable by researchers and practitioners, would increase the frequency and sophistication with which educational statistics were analyzed. The more analysis that is carried out, the more will be known about the strengths and weaknesses of the data collected. Such knowledge would result in improvements in the quality of the data collected in future studies.

SO, WHO WILL FUND ALL OF THIS?

What I have called for in this paper could easily be dismissed as being beyond the resources of the Department of Education and therefore unfeasible. Surely, anyone who suggests more federal activity these days would seem to be out of touch. There are two general ways to respond to the legitimate concern about new expenditures. The first of these is to do more by making better use of current resources. The second is to convince policymakers that spending money on statistics will not only improve learning but will facilitate more cost effective policy decisions. Let me consider the second of these arguments first.

It is worth noting that the Office of Management and Budget (OMB) often "discourages" statistics gathering proposals on the grounds that they lack policy relevance. What is being proposed here is to increase the relevance of statistics to policy making.

All levels of government expend about $120 billion on public elementary and secondary education. The NCES budget is an infinitesimal proportion of this and if the costs of all other education data gathering efforts were added up, the relative level of investment in statistics would still be minuscule. If the information developed from the types of school improvement-related data I have urged be collected were to better inform the decisions of one large school system, the cost of the national effort might be covered. Among the types of decisions that might yield large dollar returns that could be informed by the sophisticated analysis of quality data of the sort discussed above are choices about class size, beginning teacher salaries, the use of economic incentives, investments in traditional staff development and other matters. Of course, analyses could point us toward higher expenditures but, if so this would result in greater productivity. If investments in schooling could be tied to post-school outcomes, the potential for return on the investment in statistics would be very great. Of course, statistical data seldom provide, by themselves, clear policy directions. But they can challenge myths, they can suggest new options, they can confirm predispositions and they can identify issues worth further investigation. Moreover, when cynics charge that statistics can be misleading and often provide incomplete pictures the answer might be: compared to what? Most policy choices are made on the basis of intuitions informed by predispositions and by the concern decisionmakers for their political and occupational futures. Good statistics can serve as a balance wheel, if not a steering wheel, in the policy process.

What about making do with the resources we have? To pursue the course suggested in the heart of this paper, it would be desirable, perhaps necessary, to conduct one study a year of the size and sophistication of a "wave" of the HSB Study. Multi-year longitudinal research does not seem necessary though, of course, that too would be valuable. One possible source of those funds is to discontinue work being done now that does not seem to go anywhere. In addition, it might be feasible to combine resources from several of the NIE centers whose missions overlap the purposes to which school improvement-relevant statistics could be put.

It might also be possible, especially in view of the recent position of the chief state school officers to use the NAEP data comparatively, to connect the study of school processes in some way to the outcomes being studied in the NAEP. This would probably require in a larger sample and increased costs to conduct the NAEP but such piggy-backing would reduce the overall expense of the school improvement study even if this study bore the increased costs of the NAEP.

If a major new study or set of studies was not possible, two other options remain. One is for NCES to take the technical lead and provide some financial incentive that would encourage separate state studies that would be conducted, in part, in accord with a common design and with common basic variables being measured. A second option would be to do those several things noted earlier in the conclusion related to making more effective use of existing data.

Where there's a will, . . .

References

Brophy, J.E. (1980). Teachers' cognitive activities and overt behaviors (Occasional Paper No. 39). E. Lansing: Michigan State University, Institute for Research on Teaching.

Brophy, J.E., & Good, T.L. (1974). Teacher—student relationships: Causes and consequences. New York: Holt, Rinehart & Wilson.

Brophy, J.E., Monias, J., & Good, T.L. (in press). "Teacher Behavior and Student Achievement," in M. Wittrock (Ed.). Handbook of Research on Teaching (3rd ed.). New York: MacMillan.

Cooke, C., Ginsburg, A., & Smith, M. (1985). The sorry state of educational statistics. Paper prepared for the U.S. Department of Education, January.

Crain, Robert, L. (1984). The quality of american high school graduates: What personnel officers say and do about it. Center for the Social Organization of Schools Report. Baltimore: The Johns Hopkins University.

Epstein, J.L. (1984). Single parents and the schools: The effect of marital status on parent and teacher evaluations (Report No. 353). Baltimore: The Johns Hopkins University, Center for Social Organization of Schools.

Evertson, C., Willis, D., & Zlotnik, M. (1985). Making a difference in educational quality through teacher education. Journal of Teacher Education, 36 (May - June) forthcoming.

Green, J.L., & Harker, J.O. (1982). Gaining access to learning: Conversational, social, and cognitive demands of group participation. In L. Cherry-Wilkinson (Ed.). Communicating in the classroom. New York: Academic Press.

Hawley, W.D., (Forthcoming). False premises, false promises: The mythological character of public discourse about education. Phi Delta Kappan."

Hawley, W.D., & Susan Rosenholtz (1984). "Good schools: what research says about improving student achievement", Peabody Journal of Education, 61, entire issue.

MacKenzie, D.E. (1983). Research for school improvement: An appraisal of some recent trends. Educational Researcher, 12 (4), 5-17.

National Consortium on Educational Excellence (1984). An Agenda for Educational Renewal: A Fiew from the Firing Line. Nashville: Peabody College of Vanderbilt University.

Northwest Regional Educational Laboratory (n.d.). Effective School Practices: A Research Synthesis. Portland, Oregon.

Peters, T.J., & Waterman, R.H., Jr. (1982). In Search of Excellence: Lessons from America's Best-run Companies. New York: Harper & Row.

Purkey, S.C., & Smith, M.S.. (1983). Effective schools—a review. Elementary School Journal, 16, 27-29.

Ravitch, D. (1983). The Troubled Crusade. New York: Basil Books.

Roth, R., & Mastain, R. (Eds.) (1984). Manual on certification. Sacramento: National Association of State Directors of Teacher Education and Certification.

Samson, G.E., Grave, M.E., Weinstein, T., & Walberg, H.J. (1984). Academic and occupational performance: A quantitative synthesis. American Educational Research Journal, 21, 311-322.

Smith, Marshall S. (1984). Educational Improvements Which Make A Difference: Thoughts About the Recent National Reports About Education. Washington, D.C.: Federation of Behavioral, Psychological and Cognitive Sciences.

Organizational Efficacy as a Research Focus for School Improvement

Richard H. Hersh
Vice President for Academic Affairs
University of New Hampshire
Durham, New Hampshire

For the past six years I have been reviewing the research literature to determine what, if anything, makes some schools and teachers more effective than others. Happily, there is emerging from such research a variety of clues which, when put together into a coherent whole, seems to make a great deal of intuitive sense. What is particularly pleasing is that different researchers, in a variety of studies, are reaching similar conclusions about effective schooling. Furthermore, these conclusions are reinforced by school teachers and administrators who bring to research programs the critical eyes of school experience. This conjunction of researchers' knowledge and professional educators' wisdom increases the face validity of the findings but is only a beginning in understanding the casual relationships required in understanding what makes an effective school.

Three powerful facts have emerged. First, people run schools. How teachers, administrators, and students behave in a school setting matters and counts heavily toward determining a school's effectiveness. Second, quality and not just quantity of effort, materials, and time is what counts. Previously measured factors such as the total number of books in the school library, dollar amount spent per child, and the average number of years of teacher experience have been shown to account for little difference between more and less effective schools. Third, the curriculum of the school, which includes both what is taught and how it is taught, is important.

ATTRIBUTES OF EFFECTIVE SCHOOLS

Table 1 lists two sets of attributes associated with most effective schools. Under the heading of "Social Organization" are listed those characteristics which pervade the school building. These attributes (Clear Academic and Social Behavior Goals; Order and Discipline; High Expectations; Teacher Efficacy; Pervasive Caring; Public Rewards and Incentives; Administrative Leadership; Community Support) help promote school-wide conditions for teaching and learning across all classrooms. In essence, these are necessary social conditions which help individual teachers and students to excel.

The second heading, "Instruction and Curriculum," subsumes those characteristics which are found in the most effective classrooms. These attributes (High Academic Learning Time; Frequent and Monitored Homework; Frequent Monitoring of Student Progress; Tightly Coupled Curriculum; Variety of Teaching Strategies; Opportunities for Student Responsibility), in the context of the previously mentioned social organization factors, help promote the classroom conditions for maximum student engagement with purposeful learning activities. Please note that the line between the two sets of conditions ("Social Organization" and Instruction and Curriculum") is not hard and fast. In fact, the two sets are overlapping and interactive, complementary and reciprocal. Clear school-wide goals, for example, may not only help generate community understanding and support, but also they allow individual teachers to assess more accurately the fit between their expectations for students, students' expectations of themselves, and the curriculum.

Table 1

ATTRIBUTES OF EFFECTIVE SCHOOLS

Social Organization	Instruction and Curriculum
Clear Academic & Social Behavior Goals	High Academic Learning Time (ALT)
Order & Discipline	Frequent and Monitored Homework
High Expectations	Frequent Monitoring of Student Progress
Teacher Efficacy	Tightly Coupled Curriculum
Pervasive Caring	Variety of Teaching Strategies
Public Rewards & Incentives	Opportunities for Student Responsibility
Administrative Leadership	
Community Support	

CUMULATIVE EFFECTS

Each of the above attributes has been identified in effective school studies. However, it is important to note that simply developing one, two, or three of these characteristics at random would not necessarily result in a more effective school. The important conclusion to be drawn from the research is that it may be the cumulative effects of these conditions that have payoff. Although no one has shown which ones or how many of the above conditions are necessary and sufficient to guarantee an effective school, observers of effective schools suggest that there is an element of synergy. That is, it seems many things have to be done at once to do one thing well. It would be folly, for instance, to believe that simply increasing teacher expectations for students would necessarily lead to increased academic learning time or teacher efficacy. But, in some combination, some quality and quantity of these attributes reach a critical mass of conditions which promote student achievement. It is this combination, this critical mass of conditions which I label "organizational efficacy" and it is this construct which I believe needs to be more thoroughly developed and investigated by the U.S. Department of Education. What I am suggesting is rather complex and will require a disposition more toward qualitative rather than a quantitative assessment of schooling. Which agency tackles this issue (e.g. The National Center for Educational Statistics; National Institute of Education; Fund for the Improvement of Post

Secondary Education) is a policy question to be determined later.

School improvement efforts over the past decades, often referred to as "reform movements", have variously emphasized; Curriculum content and pedagogy, e.g., "new" science, math, and social studies; Back to Basics; Technology, e.g., television, calculators, and computers. We see a combination of these foci emerging in the school effectiveness literature and the addition of the more recent concern for selection, training and retention of teachers. I think the schooling effectiveness research provides a new lens through which to view the problem of educational improvement and it is in the context of that literature that I believe the concept of organizational efficacy resides and must be tapped if we are to move to a more complex understanding of improving schools.

Each of the attributes above is by itself worthy of serious consideration and each has been treated separately during the past decades. But what I'm suggesting here is that it is the interaction of all of these ingredients, the cumulative effect, the synergy created by the interactions that will determine if a school improvement effort results in significant change in student learning. To put it in more concrete terms, I am suggesting, for example, that we could triple teacher salaries tomorrow (which we might want to do on moral grounds), yet no increase in student achievement would occur. Or, if we required teachers to have Ph.D.'s in subject matter content as a substitute for present certification requirements, and these teachers had to face the same school conditions they face now, it is doubtful that student achievement would increase.

I am talking about more than the culture of the school or its ethos, although these are important factors. I am also talking about a school's capacity to not only change itself once (its values, expectations, standards, use of time, curriculum, etc.) but to change constantly, as a condition of organizational existence. A school organization which is never totally satisfied with itself will strive to improve continuously and will create an assessment system which allows its personnel to not only monitor student learning but also monitor the organization's own capacity to change. An organization in such a continuous improvement mode is like a spinning top-the gyroscopic force of motion is itself a form of stability. This dynamic aspect of an effective school organization is what Bruce Joyce and I refer to as "homeostasis of change" in our recent book The Structure of School Improvement.

Organizational life generates homeostatic forces, that is, forces that tend to stabilize patterns of behavior and keep them within a normative range. Homeostatic forces are similar to those physiological mechanisms in the human organism which keep life support functions within a normal range. In the social domain homeostatic forces resist attempts at change, precisely because it is their function to prevent changes that might endanger some essential aspect of life in the institution. Learning to live within any organization generates homeostatic pressure, and schools are no exception. The cellular model of the school in which individual teachers hold sway over particular domains and functions (the third grade, sophomore English, etc.) has greatly affected the school's receptiveness to initiatives for improvement. Inside these cells there is considerable autonomy. Most teachers work in relative isolation, with almost total operational authority over the domains to which they have been assigned. Administrative coordination in most schools emphasizes management matters such as attendance, record keeping, transportation scheduling, the cafeteria, and disciplining specific children, with much less attention to curriculum and instruction. Teachers overtly complain about their isolation but nonetheless often struggle to maintain it because, within their domain, roles are well defined and outside there is a very unpredictable milieu. They prefer social interchange which does not directly challenge their functioning.

In most schools there is a tacit understanding between administrators and teachers that their respective domains are not to be encroached on. Informal sanctions are applied to individuals who violate the norms of privacy in the classroom, or attempt to generate systematic change that affects working conditions. Teachers apply social pressure to principals to avoid direct supervision and possible changes. Similarly, principals expect to manage logistics and community relationships with a minimum of collective decision making.

Homeostatic forces are brought into play when "change agents" enter the system. The chief homeostatic mechanism is informal social pressure. Teachers in most schools are not well organized in terms of formal collective decision making but they are very well organized in terms of generating negative social pressure. For example, if principals wish to visit their classroom and offer clinical support, teachers' primary mode of counterattack is to disparage them and suggest that they are not competent to carry out clinical functions. If a curriculum change is initiated by central office personnel, resisting teachers dismiss it as "theoretical nonsense." University professors are regarded by many as "uselessly abstract" and innovators as "faddists." Disparagement is not reserved solely for outsiders who would bring innovation into the scene but is also directed toward other teachers daring enough to innovate. In many schools the innovative teachers have become social isolates.

The combination of autonomy in the classroom, relative lack of formal structures for decision making, low levels of supervision, and the use of informal social pressure to maintain classroom privacy and resist collective decision-making, result in paradoxical findings regarding teachers' feeling of efficacy. Surveys report that many teachers believe they have great autonomy within the classroom but are powerless with respect to overall district and schoolwide decision making. In most schools and district teachers are relatively powerful within the classroom but experience low levels of involvement at other levels of the organization. The isolation of the classroom increases teachers' power within it while reducing their power outside it.

This duality presents a serious problem because the effective school research suugests that school improvement requires collective activity. Any attempt to create a better environment for education will have to decrease isolation, increase cooperative planning, and sharply lengthen the amount of time in meetings. There will be a corresponding lessening of the autonomy of the classroom and an increase in teachers' efficacy in schoolwide and district planning. Unless collective activity becomes the norm homeostatic forces reign and the move toward increased organizational efficacy is stifled. Because homeostatic forces are usually more powerful than innovative forces at every level of education, ad hoc structures have to be created to promote innovation and to protect against homeostatic forces. In the absence of an executive role that promotes innovation, the necessary conditions (collective ownership, marshaling of resources, development of training, and community involvement) have to be created each time a decision to innovate is made and these conditions have to be sustained if the innovation is to persist.

To eliminate the need for ad hoc executive and protective authority calls for a substantial organizational change, one that permits reasonable and continuous innovation to take place. The condition that must be created is a homeostasis of change, a condition in which organizational stability actually depends on the continuous process of school improvement. Innovations, occasionally large but mainly small and practitioner induced, need to be normalized. To make this form of organizational efficacy happen is no small order and there are no "five easy steps" to success. Organizational efficacy occurs, as I suggested above, when a school attains both a particular level of excellence in each of the above attributes and the ability to improve continuously.

How to Assess School Organizational Efficacy

Organizational efficacy is obviously linked to specified outcomes. Since a school's purpose is multi-facted, ranging from basic skills, to critical thinking, to citizenship skills and values, these purposes will have to be carefully articulated and criteria for assessment specified. But that is precisely the function of the "clearly stated goals and purposes" in the schooling effectiveness literature. And, we need to continue to debate the issue of schooling priorities in a technological society, but that is not the purpose of this paper.

The quantitative data already being gathered by a variety of agencies, including the National Center for Education Statistics, is a necessary but not sufficient condition for the assessment of organizational efficacy. While standardized student achievement measures across schools tend to be minimalist, they provide a beginning, common basis for comparison of effects across schools, whose purposes and student populations are avowedly similar but whose effects are different. Such data, however, tend toward basic skills and rarely, if ever, tap the higher order learning involving analytical thinking, for example, which we hear is increasingly important in a world of ubiquitous data. What we need in addition, are hundreds of in-depth case studies such that each study can help illuminate the meaning of organizational efficacy for a particular school and help us generalize to that level of a critical mass of attributes needed under different conditions to achieve specific purposes for any school.

The ways in which the "effectiveness" variables work in schools is not easily quantifiable. Up to now it has been difficult to assess to what degree administrative leadership, for example, accounts for a school's efficacy versus, let's say, high expectations or a tightly coupled curriculum. Likewise, there has been no way to tell whether good instructional practice can compensate for poor materials or good materials for poor instruction. In short, the relative importance of each variable is unknown.

The search for the degree to which each variable is applicable and how it contributes to effective schooling ought to go on. I suggest that a concurrent research effort should concern itself with the more synergistic impact of the collective set of effective school variables, a set I am labeling organizational efficacy. Such an effort would require hundreds of case studies rather than the use of standardized tests. Several years ago Tomlinson pointed out that we should take comfort from the emerging evidence:

> It signifies a situation we can alter. The common thread of meaning of all that research has disclosed tells us that academically effective schools are "merely" schools organized on behalf of the consistent and undeviating pursuit of learning. The parties to the enterprise-principals, teachers, parents, and by fait accompli, students-coalesce on the purpose, justification and methods of schooling. Their common energies are spent on teaching and learning in a systematic fashion. They are serious about, even dedicated to, the proposition that children can and shall learn in schools. No special treatment and no magic, just the provision of the necessary conditions for learning.

Focusing on "just the provision of the necessary conditions for learning" is to focus on organizational efficacy. To do so we will need "thick" descriptions of school reality which only case studies provide.

An example here may help. Several years ago I visited the North Carolina High School of Math and Science. There one sees a small residential public school with all of the effective school attributes in place and a selected group of students with high motivation and high past achievement. In sum, the conditions of teaching and learning in this school are optimal and unique to public schools - small classes, fewer classes to teach, extensive teacher preparation time, adult excitment, opportunities to work with individual students, etc. The fact that this school achieves so well is not surprising but it is also not by chance. The monetary and time resources, the commitment of personnel, and the willingness to improve constantly, all combine to create an organizational efficacy which can, I think, be explained as a contributor beyond the fact of having selected the best students in the state. Indeed, the students, who are juniors when they enter, testify to extreme differences in the comparison between their old high schools and this one, exclaiming that "I never knew how much there is worth learning" and "I never knew how much _I_ could learn."

What an assessment of organizational efficacy can do through the case-study method would be to inform us of what school conditions _together_ seem to explain significant and pervasive student achievement, not to mention student and faculty satisfaction. Nor should this type of study be limited to public and private K-12 schools for a great deal of schooling is now being carried out by private and sometimes federally funded job training centers whose organizational efficacy too can be made more effective.

THE POLITICS OF EDUCATIONAL DATA COLLECTION

Paul T. Hill

June 1985

Prepared for the National Center for Education Statistics

INTRODUCTION

This paper is about the politics of educational data collection. It will discuss the human side of data collection, rather than the technical side. It will explore the sources of resentment that can lead state and local education agencies to evade federal data requests or provide inaccurate responses.

My choice of this topic was stimulated by "The Sorry State of Education Statistics," by Cooke, Ginsburg, and Smith. That paper's great contribution is to call attention to the importance of deliberate misreporting in determining the quality of federal education data. Technical improvements, such as those emphasized in NCES' call for papers, are very important, but they get at only part of the problem. The part of the problem that is not amenable to technical solution is the respondent's interest in cooperating with the data collection effort, and ensuring that his or her answers are accurate.

In this paper I draw primarily on my own experience: first as the director of a major federal data collection effort (the NIE Compensatory Education Study, 1974 -1977); and second as the principal investigator on a number of Rand studies about state and local responses to federal education requirements. In the first role I had to negotiate with state and local education to gain their cooperation with NIE's data collection efforts; in the second I expressly set out to understand the causes and consequences of the strong antagonisms between federal and state officials that I had observed during the NIE study. Both experiences convinced me that much of the political backlash against federal education programs in the late 1970s was ultimately founded in personal and professional rivalries between regulators and the regulated. Those rivalries were inevitable because federal agencies tried to impose constraints in areas that state and local officials thought were their own business.

Though I thought that the polemical atmosphere that prevailed in the late 1970s was harmful, my goal was not to assign blame for it. The atmosphere was simply the result of a sustained rivalry among federal, state, and local administrators. The natural human tendencies to form unfavorable stereotypes of rivals and to seek outside support by delivering lurid accounts of rivals' misdeeds helped to heat up the conflict. My goal was to understand the roots of the conflict so that its negative effects on educational politics and management can be controlled.

In the case of national education data, it was clear that the conflict reduced the quality, standardization, and timeliness of state and local reporting. State and local officials often understood the goals of federal data collection efforts in ways that might astound federal officials. It was therefore inevitable that many would comply minimally with requests and make serious efforts to avoid federal impositions on their time and independence.

This paper reviews the factors that lead state and local agency officials to resist federal data collection efforts or provide low-quality responses. It then identifies some potential correctives -- federal government actions that might make it easier for states and localities to understand and cooperate with national data collection programs.

SOURCES OF STATE AND LOCAL RESISTANCE

Administrative Burden

This is the most frequently-cited cause of resistance and resentment against federal data collection efforts. It is a good rationale for state and local resistance because it is easy for officials to articulate, and it makes perfect sense to members of the public who also feel harrassed by federal reporting requirements. But burden is not just a rallying cry: it is a real problem. In the course of my Rand research, it became clear that federal data requests are seldom treated as part of the routine organized work of state and local agencies. Except in the largest and best-organized school districts, they are additional loads that intrude on the schedules of already fully-programmed staff members. State and local agencies can organize their work to make federal requests seem less burdensome, but they have little incentive to do so. The onus of ameliorating the "burden" problem consequently falls on the sponsors of federal data collection efforts.

NCES's recent efforts to reduce the number and complexity of requests strike at the heart of the problem. But burden will remain an issue as long as state and local administrators believe that federal agencies make assignments cavalierly, without careful assessment of the need for data or of states' and localities' real obligation to provide it. That leads to the second source of resistance.

Federal Presumptiousness

In strictly legal terms, states and LEAs accepted the responsibility to answer federal data requests when they first took federal grant funds in the mid-60s. But many of the obligations were imposed post-hoc, and bear little obvious connection to the administration and evaluation of today's grant programs. Furthermore, many of today's administrators were not around when the orignial contract was made, and either do not know about it or feel no personal obligation to live up to it. Thus, in today's context, the simple assertion that the locals have a legal obligation to provide data is not very effective. Respondents need to be convinced that the data are going to be used for a plausibly important purpose, not simply to sustain a federal bureaucratic routine. Some possible ways to help make the purposes of federal data efforts are discussed below.

Fear of Harm

Overt opposition is not the only form of resistance to federal data requests. Many agencies are afraid to ignore requests, but resist by providing flawed or incomplete information. There are two basic motives for such resistance: the desire to avoid enforcement actions, and the wish to avoid embarassment at home.

Avoiding Enforcement Actions. Local officials know that some federal agencies gather data that can trigger compliance reviews or be used to frame lawsuits. Though officials in the more sophisticated school districts know the difference between NCES' (Or NIE's) data collection and, say, OCR's, officials in smaller districts often do not. To most local administrators the federal education bureaucracy is a big black box. Distinctions that seem utterly clear in Washington --

between audits and sample surveys, and between OCR compliance reviews and exploratory research -- are not at all clear to many local administrators. In the course of my research for Rand it became clear that local officials would regard my colleagues and me as potential informants for federal enforcement agencies until we proved otherwise. They routinely assumed that any data collection effort was the most threatening kind imaginable.

Avoiding Embarassment. Local officials are understandably reluctant to give NCES data that could make them look bad. This motive is especially intense when the data are or could be used in inter-state or inter-district comparisons. But it applies even when the study sponsors have no plans to identify the agencies from which the data were collected.

The more sophisticated school districts are not unwilling to make public disclosures, even of sensitive budgetary and student performance data. They often devote considerable resources to collection, analysis, and publication of just such data. Given the degree of public scrutiny such reports get, most are very careful to maintain decent professional standards of analysis: bad news gets reported carefully, but it gets reported. Such agencies are doubly reluctant to give raw data about themselves to anyone else. Others may not adhere to as high a professional standard of analysis as the district's own research of evaluation division maintains; and whatever the quality of analysis, outsiders (including federal agencies) are unlikely to be as careful as local officials about about the timing phrasing of disclosures about the district's problems and accomplishments.

The avoidance of inter-district comparisons may be a less important motive now than in the past. The public and elected officials now expect such comparisons to be made, and are not likely to support educational administrators' efforts to withold data. The wide attention given Secretary Bell's interstate comparison chart, the continuing strength of the accountability movement, and th4 AFT's and NEA's new acceptance of testing and comparisons among teachers all reduce the legitimacy of LEA efforts to withold data.

HOW TO MANAGE OR REDUCE RESISTANCE

The following suggestions are arranged to correspond roughly with the foregoing list of problems. But the correspondence is only rough: no one recommendation is a perfect solution to any one of the problems. But taken together the list of recommended actions will, I think, greatly reduce the severity of state and local resistance.

Reduce Data Reporting Burden by Avoiding Universal Surveys

The premise of this recommendation is that school districts will resist federal data requests less if they get fewer of them. The complaints against federal data burden could be significantly reduced by a greater use of sample surveys. Although samples undoubtedly produce less precise estimates than population surveys, they are likely to produce better data in the long run. If each national survey involves only a fraction of the LEAS, the number of data requests to a given LEA can be reduced. This will particularly benefit the smaller school systems that are the least well equipped to supply data. They are likely to fall into sampling strata that have many members, and will therefore rarely be chosen for a study sample. Because larger districts usually fall into sampling strata that have relatively few members, they will be chosen more frequently than smaller districts. But the larger districts will still experience some reduction in their response burdens.

Make Greater Use of Contractors to Collect Data

My conflict of interest is obvious here, but I will make the point because I think it is true: contractors can usually get better respondent cooperation than federal agencies can. The reasons are simple: local educators can more readily believe that professional research firms are interested in doing research, not compliance reviews. Second, individual research firms can build reputations for fairness and professionalism that put respondents at ease. That is very difficult for any government unit to match, for reasons discussed above: the differences between government agencies that do research and those that do investigations or enforcement are not readily apparent outside metropolitan Washington.

Contractors may not be able to substitute for NCES in the conduct of mandatory surveys of the entire national LEA population. The success of those studies probably requires the implicit threat of a tangle with the Department of Education for non-cooperators. But private firms are likely to get far better cooperation -- and results -- for the smaller sample surveys and exploratory studies.

Report Study Results to the Participating LEAS

School systems will contribute more willingly to NCES studies if they expect to benefit directly from the results. The largest and richest LEAS run their own data programs, and may see little need for supplements from NCES. But the vast majority of school systems could use more data, particularly about student and teacher characteristics, than they are able to collect or analyze themselves. NCES data collection would be more welcome if local officials knew that it would ultimately produce information they could use to do their jobs better. Of particular value would be information that LEA officials could use in reports to their own school boards and the public. If local officials knew that NCES data collection lead to the creation of such reports, their resistance to it would surely be much reduced.

Because many districts lack the machinery and analytical talent necessary to use raw data, this suggestion implies special work on NCES' part, to report the data in forms that school districts themselves want to use. NCES should offer participating school districts a menu of possible reports that could be created from the data being collected. These reports could be simple tabulations and non-inferential statistics that might be supplied with brief interpretive texts. Preparing such a menu would require a rudimentary market survey by NCES, to identify the range of alternatives that LEAS would find useful. It would then be necessary for NCES to build analytical routines to that could produce any of the reports automatically. A small special NCES staff would be required, to perform quality control and continually monitor the adequacy of the menu. The reports to LEAs should be data-driven and non-inferential, so there should be no need for complex text-writing, editing, or clearance.

Make Allies of Members of Congress

An important way for NCES to ensure cooperation from state and local education agencies is to make allies of Members of Congress. Members of Congress are naturally sympathetic toward public agencies in their constituencies, and ready to support them in disputes with federal bureaucracies. State and local officials who ignore federal demands -- whether substantive rules like civil rights regulations or procedural requirements like data requests -- can expect their Senators and Representatives to help if the going gets rough. This is especially true if the requirements are based on subtle administrative rationales or reflect political agendas that the Members of Congress do not support.

The lack of positive Congressional support has been a major problem for many educational data collection efforts. NCES, the evaluation divisions of USOE and the Department, and NIE have all run afoul of LEAs that refuse full participation in studies, and either threaten or actually do pull their Congressional representatives into the dispute.

But Congressional support for state and local resistance is not universal or automatic. When the political imperative behind a requirement is obvious or when the relevant federal activity is clearly useful and productive, Members of Congress are unlikely to support their constituents' refusal to cooperate.

The best way to reduce Congressional support for local agencies' refusal to provide data is to make the value of the data collection effort evident to Members of Congress. The recent use of NCES data in widely-discussed national reports on teaching has increased support among Members of Congressional education committees. But to gain support among the majority of Members of Congress it is necessary to provide information that is directly relevant to the individual Members' consitiuencies.

An annual report on the status of education in each state and Congressional district would be a good demonstration of the value of NCES' efforts. The reports should not entail new data collection; the financial, administrative, and student data that NCES now collects should be sufficient for most Members' needs. The key is to focus the reports directly on the Members' own constituencies and to deliver them directly and with some fanfare to the Members' offices. The design of such reports can be refined over time, as Members express their interest in specific information and modes of reporting. But the most important gain for NCES will be registered quickly, as Congressmen, Senators, and their staffs come to recognize that federal education data collection efforts help them understand their own constituencies.

Negotiate with OCR to Reduce Their Data Demands

OCR's school district surveys are a real problem for NCES. School systems properly count the OCR surveys as part of the overall federal data burden; and their fear that data requests can lead to enforcement actions is largely based on OCR's use of survey results. In these ways, the OCR data program handicaps NCES studies. Most of these negative effects could be avoided. OCR could conduct sample surveys, imposing data burdens on only a small fraction of LEAS, without hurting the quality of its data or reducing itss ability to target for compliance reviews: even a small sample could identify more places with suspect patterns than OCR could ever investigate.

Seek Advice From CSSOs Individually, Not in Groups

Few researchers have difficulty gaining the Chief State School Officer's approval for data collection in a particular state. As individuals, Chiefs generally have a broad policy perspective and are eager to cooperate in studies that might illuminate important national issues. Their cooperation is not automatic -- they need a good explanation of the study's importance -- but they usually answer a request quickly, without invoking complicated procedures.

Dealing with Chiefs in groups or through organizations can be a very different story. When a data collector seeks clearance from a group he or she confronts the tendency for the entire group to support individual members' objections. Each member may have one or two objections that could be readily resolved in direct negotiations. But if the group aggregates individual members' objections, its collective judgment may be that the study's problems are insuperable.

"Clearance" groups have a particularly fierce dynamic: they must pass a negative judgment every now and then to maintain morale and prove to constituents that they are doing a job. This applies to federal forms clearance organizations as well as to external groups representing SEAs or LEAs. In dealing with such groups I am frequently reminded of an instruction that Franklin D. Roosevelt reportedly gave to his staff: "Find me a bill to veto: I want Congress to know that I'm still here."

The most difficult forum for the clearance of education research plans is an advisory group composed of mid-level representatives of SEAs or LEAs. The indivudual members of such groups are serious and competent, but they have little to gain and something to lose from approving a data collection request. No one will ever thank them for clearing a study that later proved to be very valuable; LEA officials are likely to complain about the data burdens imposed by a study, whatever its ultimate value.

As individuals, the Chief State School Officers are best equipped to weight the likely burdens of a study against its ultimate payoff. They can and will complain about undue burdens and will require data collectors to accommodate the needs of local administrators. But they also can and will support a study that is needed and well designed.

CONCLUSION

The foregoing suggestions are not guaranteed to eliminate state and local resistance to NCES data collection. But they should certainly reduce it. There is, however, a cost. To reduce resistance NCES must invest the staff time and other resources necessary to design more parsimonious sample surveys, assess the needs of LEA officials and Members of Congress, negotiate with other federal agencies, and deal with important stakeholders like the Chiefs individually, rather than in groups. These are major costs to pay. NCES professionals would probably prefer to spend their time improving data definitions and analysis routines. But politics is time-consuming, and state and local resistance is a political problem, not a technical one.

INFORMATION FOR EXCELLENCE AND EQUITY IN EDUCATION

Asa G. Hilliard III

The National Center for Educational Statistics (NCES) has served a vital function for educators and planners. This has been true in spite of the limitations that mar present data gathering efforts. Cooke, Ginsburg and Smith (1965) have written in some detail about what they refer to as the "Sorry State of Educational Statistics." The state of educational statistics is most likely a reflection of certain historical policy orientations more than any deficiency of a technical order.

Americans have had a long and continuing struggle over the place of public education in the nation. (Hilliard, 1984) Should free education be provided to all citizens from kindergarten through twelfth grade? Is there a National role in education and, if so, what is it? We have lived with a system where the ideology of local control of education has been predominant. Initially, both state and national involvement were viewed with reservation. An yet there has been a steady drift toward more and more centralization of support for and centralization of control of education at the state and national levels. An so we find ourselves with a historical tradition of local autonomy and with a growing central tendency toward centralization of support and control. This affects our new data collection needs.

State and national level policymakers and leaders need to have information in order to exercise their functions. Therefore, it is necessary to continue to adjust the data gathering system so as to produce appropriate information. That information must be accurate, reliable, comprehensive, timely, representative, meaningful, and useful.

Recently, a chart was published by Secretary of Education Terrence Bell. It compared states on various dimensions. The chart triggered one of the most vigorous reactions to date by educators and non-educators alike.

Secretary Bell's chart has accomplished one thing, if nothing else. In an effort to defend themselves and to explain lower than expected state rankings, some Chief State School Officers and others who are sympathetic to their plight have been forced to articulate and to publicize critiques of the system which might otherwise have been heard only by a few. Certainly, the level of debate on these matters has been escalated and that is good. The problem remains, however, how do we take the opportunity presented by such escalations in the level of debate to improve our practices for the benefit of the children.

As has already happened on at least a few occasions, chief state school officers have taken the initiative to clarify and to standardize some practices where possible and appropriate. In the absence of such successful collaboration, the effort to develop valid, reliable, and useful information for national and local policy planning will be a waste.

It is time that we accepted, once and for all, the fact that education of children in the nation is a public matter, just as is the health of the nation. Whether the health services are publicly supported or privately supported, we recognize a clear public interest that requires public oversight. Education is no less a priority. Whether education is supported publicly or privately, the education of all children is in the interest of state and national government levels and to the public at large. This justifies major efforts such as the current effort to develop longitudinal and cross-sectional data for planning and evaluation purposes.

I see no serious problem with current data gathering *categories* at the NCES. What is needed is more refinement and expansion of existing data gathering services. I will attempt to address this latter point. It is the refinement and expansion of the focus and scope of present series that will improve the technical quality and utility of NCES data programs.

I have chosen to group my responses and recommendations into two general but overlapping categories, *excellence* and *equity*. In general, those data that support our ability to move toward excellence are also data that support our move toward equity. The reverse is also true.

I believe that we want a sytem of education that serves all children well. To reach that goal, we need a clear picture of what is going on in the schools.

Clearly, the efforts of the National Center for Educational Statistics is a macro effort. It can serve some needs. Other efforts, research, and site visits, for example, are required to round out the picture. The efforts of the NCES should be evaluated against our requirement for *general* information.

Equality of Educational Opportunity

While it is unlikely that general inequity in society can be eliminated or reduced significantly by the activities of educators alone, at the very least, educators must struggle to eliminate inequities in *educational* opportunity. This requires that areas of inequity or potential inequity in schools be illuminated and examined on a regular basis.

Traditional areas where inequities appear to occur in school settings include such things as differential drop-out rates among groups of students; high transiency rates among teachers and students; differentials in the distribution of teachers in assignments by teacher preparation and experience; differentials in expenditures per child, etc. There is a general absence of information about the wide range of diversity in the treatment of children. Therefore, when academic achievement results are very low for some groups of children, some educators have failed to examine variation within the system of treatment itself in order to pinpoint inequity. Instead, they have engaged in what Ysseldyke and others (1982) called "a search for pathology" in the children as individuals, or even within ethnic or cultural groups of children. A National Academy of Sciences Panel (Heller, Holtzman, and Messick, 1982) has suggested a different strategy. When children fail to perform, there should be, at first, an attempt to rule out the effects of what could be a low quality of educational treatment.

Naturally, no gross national data gathering effort can provide diagnostic information for an individual child or school site in order to design remedial work. On the other hand, at a macro level, it may be possible to spot situations that call for closer examination. For example, if it is shown that teachers who have the greatest amount of academic work in mathematics at the college level are not likely to be assigned to work in low income, poverty areas—this would be a situation that would signal the need for closer scrutiny.

A refinement in data collection indices may provide the possibility for isolating more accurately the effects of educational treatment on students as contrasted with the effects of certain non-school factors. In order to be in a position to address policy issues from an equity perspective more appropriately, the following types of data should be collected.

A. Equate public school and private school data collection. To the extent possible, the same types of data should be collected for both public and private schools. At present, much of the public school data are census data whereas virtually all of the private school data are sample data. Given the wide diversity of types and quality among private schools, there is some question regarding the extent to which small sample of these private schools can be considered to be nationally representative. For example, some private schools maintain very high quality systems pre-K through 12th grade. They begin with what some have described as "college preparatory kindergartens" for a student population that remains relatively stable as well as homogeneous, ethnically, and economically, throughout the full elementary and secondary school period. Other private schools are hardly selective at all. They may also offer a much poorer quality of instruction. There is a need to be able to identify such wide variations in treatment among private schools. Clearly, children vary in terms of the quality of educational experiences to which they have been exposed. By collecting more complete data from private schools, more extensive analyses will become possible. It is not a matter of collecting different data so much as a more intensive data collection effort expanded among private schools.

Recently, much ado has been made over the relative quality of achievement for public school students as compared to private school students. Yet, few data exist that help describe the types of treatment offered to students in the two types of schools. As a result, some analysts have suggested that low capacity students attend public schools.

Specific Recommendation: As much as possible, collect the same data from private schools as from public schools.

B. Level and type of non-school support. Many children in the nation are privileged to have special types of support for their academic growth from nonschool sources. These things are seldom taken into account in the evaluation of strengths and weaknesses in schools, especially the public schools. They are seldom taken into account in the evaluation of student efforts. Yet, any appropriate interpretation of statistics that are collected should be based upon the most accurate information possible. It is especially important to know the actual starting point for individuals and groups in the schools. For example, many parents are able to provide paid tutorials to supplement the public or private school education of their children. The proportion of students who receive such assistance may be very high in some schools and may be nonexistent in others. Such inequities in non-school support cause confounding when interpretations are attempted using data on school effects. Here is another example. Many educators are becoming aware of the rapidly growing gap between students who have access to computers at home and those who do not. Such gaps may also occur between schools that serve poor children and those that serve the affluent. One would expect the effects of the gap to be manifest in such areas as computer literacy, in academic achievement (when computers are used as instructional aids), and in access to word processing capabilities for composition and paper writing. It is important to know the extent to which the use of school-related technologies results in advantages or disadvantages for students who do not have access to them. Accordingly, it is important that the National Center for Educational Statistics collect data on non-school support for academic instruction such as paid tutorials and data processing.

Specific Recommendation: Collect data on access to data processing equipment for computing, word processing, and instructional software. Collect data on amount and type of paid or unpaid after school tutorial or enrichment services.

C. Stability of teacher and pupil populations. Variation occurs in the mobility of teachers and pupils at given school sites. I visited a school recently where students in a sixth grade class were working with the fourth math teacher for the year, even though the school year was only about one half completed. In some schools there are unusually large numbers of migrant or transient children. The meaning of other data such as achievement test scores is affected by such mobility. As a result, it is important for the National Center to collect such data as can give a fair indication of the level of mobility among teachers and students.

Specific Recommendation: Develop indices of mobility for teachers, students, and line site administrators. Collect data on mobility regularly as a part of the census or sampling effort.

D. Access to pre-school. The general weight of professional opinion is that pre-school is highly beneficial for children, at least in terms of preparation for academic success in school. The High Scope Foundation's longitudinal study (____) of the effects of two years of Head Start helped to extend our concept of the benefits of pre-school education to the area of social competence. In other words, not only did the High Scope Foundation study find that later public school academic achievement was higher among children from Head Start Programs than from non-Head Start children but that their social adjustment was better. And among Head Start children who were observed after the point of high school graduation, the academic achievement of pre-schoolers was higher than fron non-Head Start students. More Head Start children were admitted to college, fewer were in trouble with the police, fewer were involved in early pregnancies, and so forth. What is important is that in spite of the near universal agreement among educators about the benefits of quality early education for pre-school for children, there are large numbers of children in America who receive no pre-school at all! Estimates show that only between one-fourth and one-fifth of the children who are eligible for Head Start are actually funded in the program. Moreover, there is wide variation in quality among private pre-school offerings, even for those children who are able to afford pre-school on their own. An appropriate assessment of elementary and secondary education requires that data be kept on the participation of the attendees in pre-school programs and, to the extent possible, data should be kept to show the amount and quality of pre-school received.

I am reminded of an experience that I had recently where five out of twelve kindergartens in a certain city were designated as "developmental kindergartens." As I spoke with educators in that school district, it was clear that, in their minds, there was almost a one-to-one correspondence between the designation "developmental" and the designation" retarded." Here was a case where children were being made to pay the price for the lack of pre-school. They were seen as retarded because of low achievement, even though they had not been given the same opportunity for early education that others had. Yet, there was no attempt on the part of school officials to account for the presence or absence of pre-school experience before designating children as retarded.

At the macro level, an analysis of achievement patterns in the primary and upper elementary grades could well be informed by data on the distribution of pre-school experiences among students. There are major public policy implications here.

Specific Recommendation: At least for the elementary school years, collect and report data on the amount and type of pre-school experience to which students have been exposed.

E. Patterns of special education placement and patterns of mobility among programs by students. Anyone who is familiar with the picture in special education over the past twenty years would have to be concerned at the shifting definitions and the variation in labeling practices caused by such definitions as populations in special need. For example, there has been an alarming growth nationally in the number of learning disabled children, apparently as a result of successful litigation challenging the validity of assessment of children in the classes for the educable mentally retarded. Yet, studies such as those by Ysseldyke and others (1982) and Glass (1983) show that there is reason to question the validity of the categories as well as the validity of treatments in special education. In order to be able to understand this picture more clearly, certain types data are needed. Among them are the following: To what extent are there "graduates" of special education programs? Is special education assignment really a one-way street, or are students beginning to be returned to regular classrooms after short interventions? Are they being served in regular classes through augmented instruction? Patterns of service in special education are beginning to become quite diverse.

Specific Recommendation: Collect and report data on the mobility of students in and out of special education, by category of service, over time.

F. International comparisons. International comparisons may be helpful in interpreting what we are doing in education and in setting the appropriate expectations for what can be accomplished in education. Often, it is easy to become lost in our own parochial environment and to see as natural things that are quite unique. For example, some of our international competitors appear not to have special education as we know it. They do not have such high numbers of children designated into such categories as educable mentally retarded and learning disabled. Some are able to provide education where the overwhelming majority of their students are able to achieve a high level of "basic skills." Their achievement floors are close to our achievement ceiling. To the extent that these comparisons are valid, they force us to raise serious questions about our estimates of what the general population students in our own nation are capable of achieving.

Specific Recommendation: Collect and report data of the performance of our students on international tests of achievement. Of special interest should be a comparison with the performance of students in industrialized nations.

G. Collecting race by sex information. During a recent study by the National Academy of Sciences (Heller, Holtzman, and Messick, 1982), it was discovered that it was not possible using available educational statistics to do analyses in order to determine certain types of disproportionate placement for children in classes for the mentally retarded. It was possible to determine if there was disproportion when comparing blacks and whites. It was also possible to determine if there was disproportion when comparing males and females. However, as an artifact of the way that data were requested and recorded , it was not possible to determine what was happening by race and sex at the same time. So, for example, the frequently reported extreme disproportionate placement of black males in classes for the mentally retarded when contrasted with other categories, could not be expressed through currently available statistics. A recommendation was made by the study panel to the Office of Civil Rights that data be collected in a way that would permit race by sex analysis. There are other areas in school experience where it will be important to be able to analyze data by race and sex. For example, there is every indication that the statistics in discipline may be like those in special education placement.

Drop-out rates, disciplinary actions, student achievement, special education placements, etc. should be reported in such a way as to enable analyses to be made both by race and sex simultaneously.

Specific Recommendation: Collect and report all student data so as to permit race by sex analyses to be performed.

H. Levels of Aggregation. A general problem with many and, perhaps most, statistics is that that the results are aggregated at a level that is far too high to permit the best analysis of what is going on. Data aggregated at the state or school district level may serve some useful purposes but, for many purposes, the most significant information is the presence or absence of a pattern of variation among school sites or even among school classrooms, sometimes within a given school site. Then, of course, as has already been recognized by National Center staff, there are times when the variation among individuals is of great interest. Given the capacity of data processing equipment to handle large amounts of data, it is important that data on most school variables be disaggregated to the lowest possible level. For example, it was not until the effective school research movement that many instances of excellence in education at regular school sites among low income schools were uncovered. For the most part, isolated schools that were "swimming upstream" were buried in aggregated data which tended to suggest that no such schools existed. In fact, analyses of much of the school effectiveness research led to the erroneous conclusion that schools had little or no effect. Questions such as "Do schools work?" were common. It is notable that follwing the effective school research, the question more often is "How do good schools work?" The same may be said of effective teacher research.

Specific Recommendation: Whereever possible, disaggregate data. Provide reports on both aggregated as as disaggregated data.

Educational Quality

During recent years, there has been an extension of the research on effective teaching and effective schools. Many of the research results have not been popularized. A few such ideas as "time-on-task," "engaged-learning-time," "locus-of-control," etc. are part of the common professional language. Yet, many of the things that have been learned from effective teaching and effective schools research cannot be used in state and national policy level discourses. The data that might suggest the need for further inquiry are not collected because of feasibility considerations. In some cases, it would be nearly impossible to collect (on a mass basis) the kind of information that is desirable, such as time-on-task by an individual student. However, there are other cases where the collection of certain data is feasible and can illuminate better the quality of the instructional offerings in the school. Every opportunity should be taken to collect this information. A few examples follow.

A. The academic preparation of teachers. The National Center is already sensitive to the problem of collecting information about teacher preparation by relying upon certification categories. Clearly, there is a lack of unity among the certification categories from state to state. The same may be said for academic majors and minors. However, it may well be that information about the academic majors and minors of teachers would be more revealing than information about the typical certification categories into which teachers fall. This information should be collected. The need for such information should be apparent. There may be equity questions involved in the assignment of teachers according to academic preparation. For instance, let us consider the areas of mathematics and science. In a large city school district that does have a full quota of certified mathematics teachers or science teachers, is there any relationship between the amount of academic preparation in mathematics and the assignment of teachers to low income and high income schools. It would be of interest also to know how the public schools compared with private schools in this regard. It is well known that some private schools emphasize academic preparation over professional preparation, preferring to hire teachers with academic majors and with academic master's degrees. Of course, this is an area where there are many, many questions. What is important is that data be available which would be useful in developing answers to some of those questions. The ease of collection of such information and the availability of national populations for study make it compelling to do so, considering the benefits which may be obtained.

Specific Recommendation: Collect and report data on the academic major and minor preparation of certified staff, disaggregated to the school site level.

B. Describing the school curriculum. Anyone who is even minimally familiar with schools is aware that there is no common nomenclature for classes that would enable a meaningful analysis to take place regarding precisely what content is offered in schools. It may well be that we are destined to be stuck with this problem in some form for quite some time. Nevertheless, it should be possible to improve upon present practices. A report such as that issued by The College Board (1984), *Academic Preparation for College: What Students Need to Know and Be Able to Do*, should be helpful in attempting to pinpoint the types of topics that may be covered in course content. It should be possible to make a compromise by collecting data that falls somewhere between the level of detail outlined in the College Board report and the gross categories that we normally use. For example, it may be very useful to know how many students have passed course work in algebra and geometry. This may be more important than knowing what the quantitative score of a group of students was on the Scholastic Aptitude Test (S.A.T). At the high school level, it is possible to identify certain key courses such as algebra, general chemistry, foreign language, first-year foreign language, college preparatory English, etc. and to determine what proportion of the students have completed the key courses. This leads us away from dependency on normative data and toward more meaningful criterion data.

Specific Recommendation: Work with Chief State School Officers to develop a common nomenclature for key academic courses. Collect and report data based upon this nonmenclature.

C. An academic success criterion. At present, the use of the S.A.T. or the A.C.T. at the end of a high school program as a measure of academic achievement is seriously problematic. The absence of meaningful and viable alternative is also seriously problematic.

There is a logic associated with the whole data collection system as it now stands. That logic may also force us inevitably to the conclusion that there is a need for some uniform measure of academic performance at the national level. To my knowledge, no test publisher has ever made claims to the effect that any instrument published by them was indeed a valid universal measure of academic achievement. Rather, users are left to determine (based upon their own analysis) if the match between test content and the academic objectives that they espouse is sufficient.

Another major problem with the use of the S.A.T. and the A.C.T. stems first and foremost from the fact that it is necessary to determine if the tests are considered to be measures of "aptitude" or measures of "achievement." Sometimes, the word "ability" is used to describe tests such as the A.C.T. or S.A.T. However, the use that is made of such tests reflects confounding in the minds of users regarding the nature of the test as either aptitude or achievement. Most often, users attempt to stand in both places at the same time—implicitly claiming that the tests are both aptitude and achievement.

The significance of this (for the National Data collection effort) is that ultimately a choice must be made between these two options. Once having been made, the tests must be evaluated according to the appropriate rules for evaluating the particular type of test that it is. For example, if it is an achievement test, it must evaluated according to the rules for determining its content validity for a high school curriculum. This brings up the awesome problem of validity of the criterion, the school curriculum. Little needs to be said about the absence of uniformity in the high school curriculum. Standardized testing for a non-standard curriculum is an absurd practice. In the absence of more uniform curricula, the test cannot be content valid.

If it is an aptitude test, then it must be evaluated according to the rules fro determining predictive validity, taking into account the variation in instructional quality that intervenes between initial teaching and final testing. For example, if these tests are regarded as aptitude tests, the results of studies of coaching effects should give real cause for pause. (Messick, 1980) (The Federal Trade Commission Study, 1979) It has been shown that standardized tests scores can be raised significantly by well-designed, short-term coaching courses. This should not be the case if the test is a test of "aptitude." Perhaps, the only reasonable resolution to this problem is to call upon Chief State School Officers to take the lead in establishing the uniformity in academic goals at a basic level that would permit test publishers to develop tests based on common understandings.

Such an approach is not without its dangers—the most obvious of which is the loss of local control over curriculum decision making. The issue here is more one of a policy matter than a technical one. Until we get to the point of considering whether certain important educational objectives can be measured by traditional forms of paper and pencil, multiple choice testing. At that point, another policy issue presents itself: What costs will educators accept for the quality assessment of academic achievement?

Many things in the national data collection plan are linked to achievement test results. Achievement test results, for better or for worse, are considered to be the "bottom line" in the data collection effort. Therefore, the stakes are very high. There is a critical need for valid outcome measures.

Specific Recommendation: Work with Chief State School Officers to develop standardization in testing that is feasible and appropriate. Collect and report achievement data from these new measures.

General

In general, the present categories of data collection are appropriate. Reliability and validity must be the major concern. This is the main way to improve present data gathering efforts.

Strong support should be given to the High School and Beyond Survey. It is one of the few places where individual students are tracked. Moreover, as a longitudinal study, it will be a rare contribution to our knowledge base.

The Library/Media Center Survey is important. However, it is not clear that qualitative judgments can be made from the quantitative data to be collected. If there were a report that summarized the holdings by titles within categories, it would be much easier to perform evaluations of the quality of holdings. For example, what is the pattern of holdings in typical schools. Summative information on these patterns is desirable.

The Twentieth Century

One of the most interesting things about progress is that the more some things change the more others stay the same. Most of us have witnessed phenomenal changes in the availability of technology such as television, computers, genetic engineering, space travel, etc. Indeed, the content of school curricula now reflect this new information. And yet the requirements for a basic elementary and high school education of quality are really not all that different today than they were decades ago. As we try to prepare children "for life" or for the "work of work," we find that both of these areas call for students who are skilled at reading, computing, analytical and synthetic thinking, written and oral expression, and a whole host of "liberal arts skills." (Adler, _____)

We should have learned by now that the best preparation for the world of work at the high school level is a good sound academic and social educational experience. First, high technology in the workplace does not seem to increase the call for "high tech" jobs. (Levine and Rumberger, 1983). The U. S. Department of Labor confirms the fact that the growth areas for employment are in the low-skilled service sectors of the economy. It is hard to train students for specific jobs that matter at the high school level. Second, the advanced level jobs and personal satisfaction in life require a sound general education, not different in kind than that which we have described many times before. (Adler, ____)

The best role of data gathering on education for public policy decision making is a role that supports the most refined description possible of what takes place in the schools. It is essentially an operation that functions in support of quality control and equity guarantees.

Ultimately, our mission in education must be to serve our people. We do that by being cognizant of the demands of the economy. But we can never neglect the thing that we have always heard articulated. A democratic society is dependent upon an educated citizenry. This means that our vision for the nation's schools is that they are instruments that build the capacity of students to think. The schools are instruments that confront them with the important things that citizens must ponder. In this regard, the twentieth century is not unlike previous centuries, except perhaps the gap betwen or ideals and reality can be closed if we can see reality more clearly.

SELECTED REFERENCES AND BIBLIOGRAPHY

College Board (1983). Academic Preparation for College: What High School Students Need to Know and Be Able to Do. New York: The College Board..

Cooke, C., Ginsburg, A. and Smith, M. (1985). "The Sorry State of Educational Statistics." Manuscript.

Federal Trade Commission (1979). Effects of Coaching on Standardized Admission Examinations. Washington, D.C.: Bureau of Consumer Protection, FTC.

Glass, Gene V. (1983) "Effectiveness of Special Education." Policy Studies Review, Vol. 2, Special No. 1. Lawrence, KS: The University of Kansas Press.

Heller, Kirby A.; , Holtzman, Wayne H.; and Messick, Samuel (Eds.)(1982). Placing Children in Special Education: Strategies for Equity. Washington, D.C.: National Academy Press.

Hilliard, Asa G. (1984). "Democratizing the Common School in A Multicultural Society." Education and Urban Society, 16, 3.

Levine, H.M. and Rumberger, R. (1983). "The Low-Skill Future of High Technology," Technology Review 86, August-September, pp. 18-21.

Messick, Samuel (1980). The Effectiveness of Coaching for the SAT: Review and Reanalysis of Research from the Fifties to the FTC (ETS RR 80-8). Princeton, NJ: Educational Testing Service.

Schweinhart, L.J. and Weikart, D.P. (1980). Young Children Grow Up: The Effects of the Perry Preschool Program on Youths Through Age 15. Ipsalati: The High Scope Press.

U. S. Department of Education (1985). State Education Statistics

Ysseldyke, James E.; Thurlow, Martha L.; Graden, Janet L.; Wesson, Caren; Deno, Stanley L.; and Algozzine, Bob (1982). "Generalizations From Five Years of Research on Assessment and Decision Making (Research Report No. 100).

EDUCATIONAL STATISTICS FOR EDUCATIONAL POLICY:
A POLITICAL ECONOMY PERSPECTIVE

Submitted to the

National Center for Education Statistics

by

Maureen McClure

University of Pittsburgh

and

David N. Plank

University of Pittsburgh

EDUCATIONAL STATISTICS FOR EDUCATIONAL POLICY: A POLITICAL ECONOMY PERSPECTIVE

The environment of the educational system is becoming more uncertain. The American economy is in a period of transition: employment in basic industries is in decline relative to employment in the service sector, and the character of the national economy is increasingly determined by an international division of labor. Birth rates are declining, while rates of inter-regional migration are high. School reform is once again a national preoccupation, and proposals for changes in basic and higher education have been put forward in virtually every state in response to perceived inadequacies in the performance of the schools. Proposals for federal tax reform threaten the traditional financial basis of the public schools, while discussion of tuition tax credits and educational vouchers challenges the privileged status of public schools within the educational system.

When faced with uncertainty, the Wicked Queen in "Snow White" ran to her mirror: omniscient, timely, and truthful, the mirror gave her the information she required. Educational policy makers have no such recourse. In uncertain times they are often obliged to act on the basis of information that is sketchy and unreliable, and of limited relevance to the decisions they must make. This need not be the case. The expanding activity of the federal government in the collection and dissemination of the data and the rapid development of the information technologies now make it possible to provide educational policy makers with access to reliable, timely data on many of the critical issues they face. The National Council on Education

Statistics should play a leading role in compiling these data and ensuring their availability to those who need them.

In this paper we discuss some of the ways in which NCES might reorient its activities in order to assist educational policy makers in making informed decisions. The paper is organized in three sections. The first section discusses the relationship between the educational system and its tax base. The second section of the paper proposes that policy relevance rather than convenience should determine the levels of aggregation of educational statistics. It also argues that decentralization in data collection will facilitate access and encourage relevance. The third is concerned with the maintenance of uniformly high standards of reliability and comparability in NCES data. It points out the dangers in confusing informational and regulatory data bases. and urges that the two be kept entirely separate.

Tracking the Investment Cycle

An increasingly complex economic and technical environment is changing the structure of educational policy making. Public support of education is at least partly grounded in the belief that education contribute to economic development. Concern about the returns to past investments in public schooling underlie many current efforts to reform the schools. The development of the future tax base must be a priority for educational policy making over the next decade. Without access to an adequate source of funding. excellent educational programs will do the next generation of taxpayers little good. The current school reform effort must recognize that education is a means to an end. and not an end in itself.

Education can be viewed as a middle product in a public investment cycle. Taxpayers can be seen as investors in those goods and services that would not receive adequate provision in completely private markets.

They invest in education primarily through property, sales and income taxes. Taxpayers in a democracy elect those who set taxes, thus exercising some degree of choice in the development of educational policies. Local, state and federal governments are then charged with investing these resources in educational programs, which in turn, are reinvested resources in students. (Students also invest their own resources in education, primarily in the form of time.) Eventually students enter a labor market and invest their human capital and time in return for a wage. They then close the investment cycle by becoming the next generation of tax-paying investors. Each investment in the sequence occurs under conditions of uncertainty, and each consequently requires an appropriate level of return to compensate for its risk. The adequacy and type of return to the educational investment is primarily a function of the investors' standards and perceptions of the returns to alternative investments.

Except at the national level, tax bases are not pooled. Local property taxes invested in a community's children may not produce a return if those children enter the labor market in other communities. If there is a balance of in and out migration, local governments, states, and regions do not have to be overly concerned about tracking their investments. If local and state investments in human capital ultimately subsidize their economic competitors, however, then policy makers will eventually restructure their educational investment policies appropriately.

The federal government can provide a service both to the education community and to national economic development by initiating a series of regional data bases that accurately track educational investment cycles in both public and private sectors. While local, state and private sources can help to support the initiative, the federal government through NCES should set standards for data collection, insure comparability and timeliness, and provide computer networking systems for decentralized access.

A high quality system of information could encourage cooperative efforts among policy makers and researchers at all levels. These efforts could result in a more flexible, more responsive, and finally more productive educational enterprise.

Education is one of the nation's most broadly-based investments in future tax base renewal. It is also a nationally contained investment. While regional economic competition may create some policy friction, from a national perspective the resources invested in citizens have only minimal leakage, because relatively little human capital migrates and subsidizes foreign economic competitors. National educational investments for development contrast with the public subsidization of private corporations for the same end. Corporations can, with relative ease, export their investments and indirectly subsidize international competitors. Educational investments present lower risks for public investors.

Unfortunately, relatively little is known about the investment linkages between education and national economic competitiveness. Even this limited knowledge is eroding in a transitional economy with a murky future. In the fifties and sixties, there was a widespread belief in the contribution of mass public education to economic development. The American model was exported to developing nations with great fervor, but with little solid information.

Challenges to this confidence in public education surfaced in the mid-seventies as the momentum of the American economy waned. These challenges have generated intense criticism of the public school system and a variety of proposals for school reform. The educational community will continue to face criticisms of its investment decisions into the next decade. Lacking adequate information, policy makers are likely to make decisions in response to the pressures of the moment, rather than to seek out productive, long-term educational investments.

According to a recent Urban Institute study, the average current wage earner from the "baby boom" generation earns ten percent less in real terms than did the average wage earner in the prior generation. Yet the "baby boom" generation was the recipient of extraordinary public educational investments of the fifties and sixties. Is this merely a temporary labor oversupply problem that will clear in the next generation, or are these returns a portent of the economy of the future? No one has an adequate mirror.

If educational policy makers at any level cannot track their educational investment cycles, they may incorrectly identify the problems requiring attention and inadvertently generate policies which result in costly unintended consequences. Educational delivery systems which may have been both appropriate and successful under different economic and technological conditions may hinder economic development under new conditions. For example, an economy with assumed growth can bear greater nonproductive distributive burdens than one with uncertain future growth.

The less that is known about current delivery systems and their impact on development, the riskier educational policy making becomes in the face of economic transition. A high quality integrated data base can provide a basis for reducing that risk.

The investment cycle can be tracked in a variety of ways. Sets of individuals should be followed through their educational careers and through the labor market, building on the High School and Beyond study.

Tracer studies should track students through alternative educational programs into the labor market. These studies should include fuller and deeper linkages with the labor markets and migration patterns.

School-site resource allocation patterns of investment can be tracked through comparable measures. NCES could carefully examine detailed expenditures across a representative sample of school sites. Comparable data could be collected on direct and indirect fixed and variable, program, and logistical support costs. Outcome measure from these sites could be used to examine the cost effectiveness of current practices. Detailed environmental data could also be integrated to the site data base.

Finally, measures of educational investments in economic development might be most appropriately tracked at a regional level, both to avoid problems of data collection at too broad a level of aggregation and to follow the flows of educational capital across regions. NCES should not engage in direct collection at this level but can play a very useful role by coordinating, and editing existing data from other sources. At regional levels these would include comparable information about tax bases and tax effort, voter responses to tax initiatives, labor market patterns and trends in employment by sector, detailed in and out migration patterns, demographic structures, economic indicators such as housing starts, and public expenditures by sector. Good mirrors are expensive, but invaluable to policy makers.

Individuals, Sites, Regions

A central issue in public data collection over the next decade will be the aggregation problems connected with public investment. There are major gaps between policy and research questions and available data. For example, how do local and regional investments in education affect the regional and national economies? What are the tax base returns from public investments in individual students? How is the school reform movement working at the level of the school site? These questions cannot be answered with present NCES data.

The aggregation problem is exacerbated by an increasing slippage between economic markets and the scope of public policy makers. Economic boundaries are fluid, as capital flows in and out of cities, regions and nations. Political boundaries are somewhat more rigid, as political stability rests on defined territory. Where economic and political boundaries are coincident, public and private interests may coincide as well. As boundaries diverge, however, sectoral slippage increases. For example, local investment in public schools may pay off when graduates remain in the community to regenerate the local tax base. If, however, the community's public school graduates leave the community and allow the returns to public investment in their education to accrue to the community's competitors, then investments in public education may result in net losses for that community.

NCES can play a crucial role in the formulation of educational policies by providing high quality information to policy makers. Reliable, valid, timely information could both provide the education sector with a competitive edge and help to maintain standards of quality in public investment. These data should be collected at natural economic levels of aggregation: 1) individual achievement and economic performance over time, 2) school site productivity and 3) regional economic returns on educational investments.

NCES in cooperation with other federal data collecting agencies should establish a national tracer data base to follow individuals through the public investment cycle. Individuals carry with them public investments in their human capital. These investments are not evenly distributed across students. Test scores measure outcomes for the relatively short term, but additional measures of intermediate and long term effects would make possible a more accurate tracking of the educational investment cycle. In addition to tests, intermediate term indicators should measure the return on the public investment in the labor market. Tracer studies should track students from different backgrounds through alternative educational programs (private/public, academic/vocational/general), into the market place. The Bureau of Labor Statistics and others could then assist in tracking their progress.

Long-term measures should go beyond income as economic indicators of a return on investment, because income does not close the investment cycle for public education. Education revenues are still drawn for the most part from property rather than income taxes, and property taxes are by definition tied to the economic health of regions. The financial basis of public education assumes stability in the investment cycle: local communities and states invest in public education so that graduates will provide returns as property owning citizens in the future. Disruptions in this cycle caused by net in or out migration disturb the community's expectations of a return on its investments in education and alter tax payers' willingness to support the schools.

Another neglected level of data collection is the school site. School sites are the most important levels of aggregation for assessing the impact of the school reform movement. They bear the burden of the regulation generated by legislative initiatives and community response. School districts, intermediate units and states are less interesting and relevant levels of aggregation than school sites for the measurement of the impact of educational policies.

NCES should select a representative sample of school sites from which to collect a rich base of information about the relationship between school site investments and educational outcomes. Individual studies have been conducted in this area, but a national longitudinal effort comparing public and private school sites across regions would provide researchers and policy makers with integrated, reliable information with which to track the progress of school reform initiatives.

Site indicators should include regulation costs to track the time and resource response costs of centralized policy initiatives. Also while NCES has collected data from private schools, a much more comprehensive picture of site costs of alternative education investments is in order.

Policy makers must wrestle with serious questions about the returns on investment in public education. Would deregulation through increased competition raise or lower teacher wages in a free market? Would deregulation increase the sector's investment in low cost labor intense practices or would there be a shift toward lower cost physical technology? Before such pressing questions can be addressed more must be known about the actual costs of educating different types of students under alternative conditions.

At the regional level, NCES can provide a valuable service by supporting integrated data bases which more closely link public investments and economic returns. Regional data bases can more accurately track the interplay between private and public investment and return; therefore, the effects of educational investments on tax bases generated by human capital circulation and savings can be measured more meaningfully.

While policy making will continue at local and state levels. regional data bases can provide decision makers with higher quality information than if data are collected at more artificial levels of economic aggregation.

The interaction between education investment and the economy is complex. and data aggregated at the national level mask the important effects of externalities generated by regional subsidization of local and other regional economies. Failure to recognize these effects may lead to inaccurate predictions about the expectations and behavior of taxpayers. and to a misunderstanding of the performance of regional economies. Over a longer term these effects could be substantial.

The federal government. by sponsoring integrated comparable regional data bases could encourage decentralized. coordinated decision making for economic development. Economic regions cut across state lines. and regional data bases would lower the costs of accurate. timely and accessible data for policy makers at local. state and federal levels by reducing the duplication of effort. increasing the data collection investment pool. and encouraging greater cooperation in regional development efforts. These data bases would not only track regional investment cycles but would also provide comparable data for tracking national trends.

Regional data bases also make technical sense. NCES can help to coordinate federal inter-agency efforts to link public investment to tax base return. The Census Bureau. the Bureau of Labor Statistics. the National Science Foundation (NSF). the Department of Commerce and many other agencies collect information which is vital to the tracking of economic development. These departments should work more closely with their counterparts at state and local levels to insure accurate. comparable. timely data at regional levels.

Recent technological developments can provide momentum in these new directions. Supercomputers can process complex dynamic models that until recently were only theoretically possible. A high quality regional data base could support more sophisticated modeling of public investment cycles.

NSF has sponsored a series of Cray supercomputers for academic use. A cooperative effort between NCES and NSF could pilot a regional data base to track a public educational investment cycle. Initially it would be better to focus resources on a single high quality data base than to lower the quality of data collection by underfunding a more broad based effort. Under these conditions. investment in a pilot data base could lead to lower-cost implementation of a refined and expanded system at a later date. State. local and private investors could be encouraged to bid for the experimental information base.

Collection Criteria

In the past NCES has not been known for high quality indicators and data collection. There should be no compromise on standards of quality in a refocused program. The criteria for NCES data collection should include: 1) parsimony. 2) accuracy. 3) comparability. 4) timeliness and. 5) accessibility.

Parsimony is paramount. Overcollection of data at low-cost levels of aggregation is wasteful and of little use to policy makers. The nature of the data collected by NCES should be determined by long-range usefulness.

What are the important questions that policy makers and researchers will be asking over the next decade? What resources are required to support long term maintenance of high quality indicators? Less is more if fewer items of data are collected at more appropriate levels of aggregation, even if this means higher costs per item of data. If NCES is to serve as a role model for a refocused federal information system, then enough care should be invested in it to make it cost effective in the long term.

Accuracy is an acute quality control issue. It requires unambiguous standards of tolerance that are frequently monitored. Indicators must be defined in clearly measurable terms. Opinion poll information can be useful to policy making if the data standards are considered excellent by experts and the methods and standards employed in collection and interpretation are easily accessible. Access to raw data for reinterpretation should be considered.

Comparability counts. Crossectional financial data comparisons are currently impossible because financial data are collected for idiosyncratic regulatory purposes. Longitudinal data comparisons have not had formal, consistent long-term tracking commitments. For example, the quality of minority representation data has eroded over the last decade. Longitudinal data comparability should require firm, ten-year commitment levels. NCES should encourage states to collect at least minimal amounts of comparable crossectional and longitudinal data, especially within regions. The lack of such data obliges policy makers to make decisions without an adequate supply of information.

The lack of timeliness in educational data collection is shameful. Lag times in reporting are often measured in years instead of days or months. This is a serious underinvestment in data collection activities. Policy makers require timely data for decision making and should be willing to pay for it.

Accessibility is an absolute requirement of a refocused federal information system. Networking systems can now be established to allow highly decentralized use of data. This requires the capacity to download files into other systems and the availability of highly sophisticated user-friendly software so that questions can be addressed with minimal inconvenience. In addition to a user-friendly data base management system, NCES should sponsor the development of expert systems to interface with central data bases so that policy makers can ask questions and receive timely answers in a useful format. Accessibility could additionally include easy interface with graphic systems, statistical packages, and "what-if" scenario packages.

Networking through NCES might also provide access through the system to data bases maintained by other agencies. Access to multiple data bases could be useful in tracking national and regional investment cycles.

High quality data collection can influence the process as well as the outcomes of policy making and research. NCES should invest in studies with NIE and other agencies to study the potential impact of this influence.

Quality can be produced if sufficient incentives are available. NCES may want to provide incentives for quality through the use of fees for services and user charges. NCES could contract with school sites and other agencies for data. Policy makers should be willing to invest in data collection that can provide them with support.

While data collected for NCES could serve to monitor regulatory costs, they should not be used for regulation for at least three reasons. First, monitoring requirements tend to result in highly specialized, complex and cumbersome data requirements. Second, regulation by definition introduces reporting biases, thus distorting accuracy and reducing the credibility of the data base. Threatened administrators could stall in reporting and processing data to reduce timeliness. Third, secrecy is often a major component of regulated systems. The NCES system should be known for its wide accessibility through multiple networking systems.

While omniscience is beyond the reach of any federal inormation system, timeless and truthfulness should be the objective of all NCES activities. NCES should strive to answer those questions that policy makers ask, not those that are easy or inexpensive to address, and every effort should be made to ensure that the answers are based on reliable, timely data. NCES should be seen by policy makers as the manager of a mirror for the educational system, and not a purveyor of poisoned apples.

REFERENCES

Becker, G.S. 1976. *The economic approach to human behavior.* Chicago: University of Chicago Press.

Benjamin,R. 1980. *The limits of politics.* Chicago: University of Chicago Press.

Bluestone, B. and B. Harrison. 1982. *The deindustrialization of America.* New York: Basic Books.

Bowles,S., D.M.Gordon and T.E.Weisskopf. 1984. *Beyond the wasteland.* Garden City, NY: Anchor Books.

Boyer, E.L. 1983. *High school.* New York: Harper & Row.

Chubb, J.G. 1981. The political economy of 'excessive' regulation. Project report no. 81-A12. Institute for Research on Educational Finance and Governance. School of Education, Stanford University.

Dean, E. 1984. *Education and economic productivity.* Cambridge,MA: Ballinger.

Denison, E.F. 1979. *Accounting for slower economic growth.* Washington, DC: Brookings Institute.

Drucker, P.F. 1980. *Managing in turbulent times.* New York : Harper and Row.

Everhart, R.B. ed. 1982. *The public school monopoly.* Cambridge, MA: Ballinger.

Feldstein, M. 1980. *The American economy in transition.* Chicago: University of Chicago Press.

Freeman, J. and M.T. Hannan. 1981. Effects of resources and enrollments on growth and decline in school districts. Program report no.81-B1. Institute for Research on Educational Finance and Governance. School of Education, Stanford University.

Freeman,R.B. 1976. *The over-educated American.* New York:Academic Press.

Garms, W.I., J.W. Guthrie, and L.C. Pierce. 1978. *School finance.* Englewood Cliffs, NJ: Prentice Hall.

Goodlad, J.I. 1984. *A place called school.* New York: McGraw Hill.

Guthrie, J.W. 1980. United States school finance policy: 1955-1980. *School finance policies and practices.* ed. J.W. Guthrie. Cambridge,MA: Ballinger.

Hannan, M.T. and J. Freeman. 1983. School district demography and governance. Project report no.83-A22. Institute for Research on Educational Finance and Governance. School of Education,Stanford University.

Kirst, M.W. 1984. The changing balance in state and local power to control education. *Phi Delta Kappan.* November:189-191.

Kirst, M.W. and W.I. Garms. 1980. The political environment of school finance policy in the 1980's. *School finance policies and practices.* ed. J.W. Guthrie. Cambridge, MA: Ballinger.

Magaziner, I.C. and R.B. Reich. 1982. Minding America's business. New York: Vintage Books.

Pogrow, S. 1983. Linking technology use to school improvement. School finance and school improvement. ed. A. Odden and L.D.Webb,127-142. Cambridge, MA: Ballinger.

Reich, R.B. 1982. The next American frontier. New York: Times Books.

Rumberger, R.W. 1984. The potential impact of technology on the skill requirements of future jobs. Project report no. 84-A24. Institute for Research on Educational Finance and Governance. School of Education, Stanford University.

Schultz, T.W. 1982. Investing in people. Berkeley: University of California Press.

Sizer, R. 1984. Horace's compromise. Boston: Houghton Mifflin.

Smith,J. and F. Welch. 1978. The overeducated American? Vol.5. Proceedings of National Academy of Education.

BEST COPY AVAILABLE

A Model for N.C.E.S. Research on School Organization and Classroom Practices

James M. McPartland

Henry Jay Becker

Center for Social Organization of Schools

Johns Hopkins University

September, 1985

This paper was commissioned by the National Center for Education Statistics as part of its effort to design a 10-year program of data collection to improve our understanding of the nation's elementary and secondary education systems.

A Model for N.C.E.S. Research on School Organization and Classroom Practices

N.C.E.S. has two principal data collection functions. The first is to produce, on a regular and recurring basis, important descriptive statistics on the functioning of educational institutions. Recurring descriptive statistics provide comparative and historical measures that help us understand how educational services are provided to people in different places and circumstances.

The second data collection function of N.C.E.S. is to provide information not otherwise available to improve our understanding of the consequences of different ways of providing schooling to students---consequences which, if better understood, could lead to the improvement of the delivery of educational services. It is this second data collection function--gathering data for understanding and improving how schools provide instruction--that is our own primary interest and is the focus of this paper.

The paper has three parts. First, we present a model of important aspects of the provision of instruction in schools that prior research suggests may make the most difference in student outcomes. Second, we describe the type of research that we feel is necessary to answer questions implicit in the model, and we outline a specific research design that addresses these issues in a concrete way. And third, we provide a selection of survey questions that would form a portion of the information needed in a research project such as we have outlined.

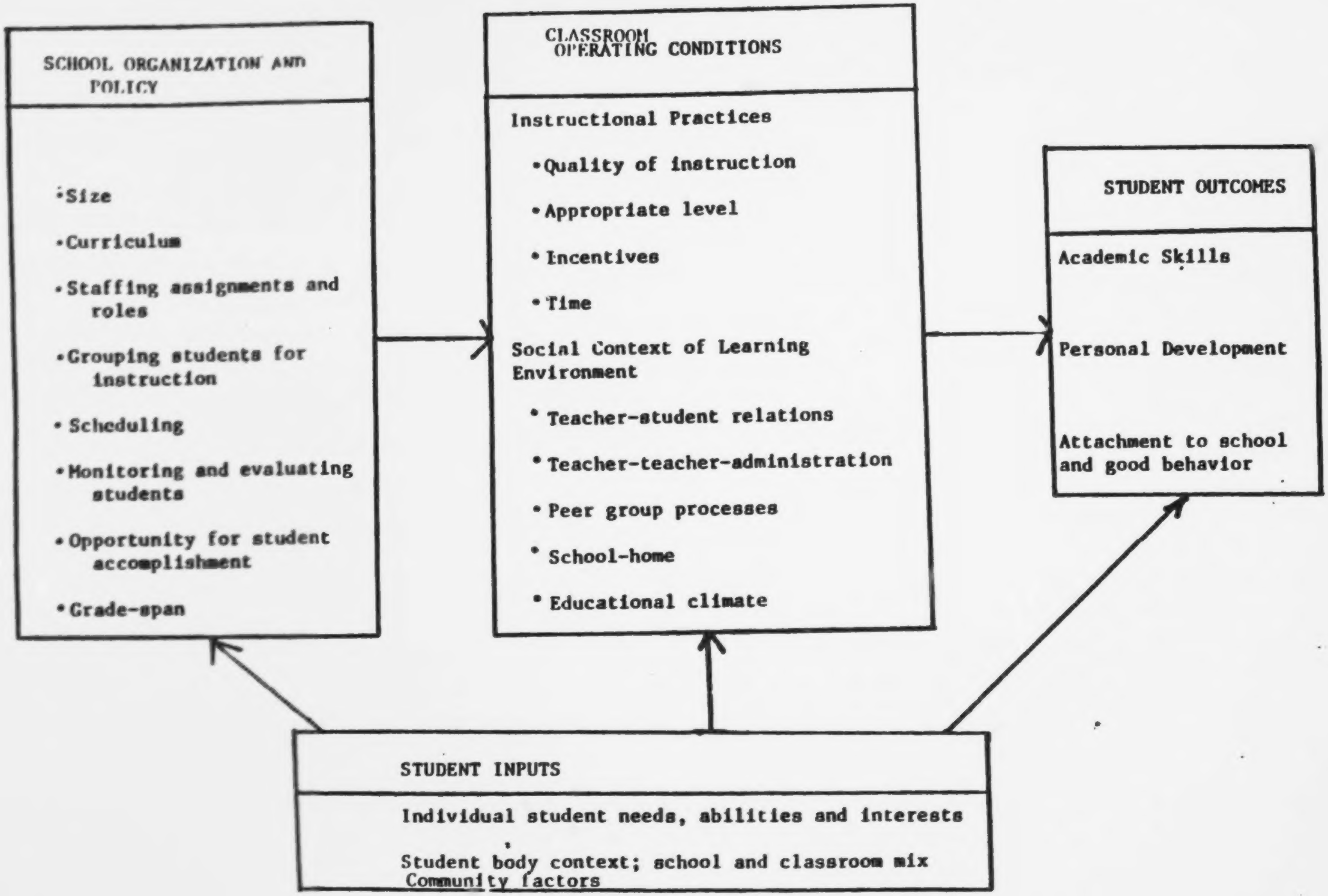
I. A Model to Guide Research

In order to appropriately discuss data collection strategies and designs for an N.C.E.S. program of providing data for school improvement, we must begin with a model of the dimensions of school instruction that are alterable by policy and training of the participants and for which alterations might be likely to improve at least some important outcomes for students.

Figure 1 presents such a model of school factors and student outcomes. The model includes two key working assumptions. First, student learning is most strongly affected by the most proximate influences on the individual: one's own home and one's own classroom. To represent this, we have separated general school factors from immediate classroom conditions in our model, and we show causal impacts on student outcomes to come primarily from classroom conditions and student inputs, which include home background.

Second, although classroom conditions have the most immediate impact on student learning and development, these classroom conditions are

Figure 1


SCHOOL ORGANIZATION AND POLICY
•Size
•Curriculum
•Staffing assignments and roles
•Grouping students for instruction
•Scheduling
•Monitoring and evaluating students
•Opportunity for student accomplishment
•Grade-span
CLASSROOM OPERATING CONDITIONS
Instructional Practices
•Quality of instruction
•Appropriate level
•Incentives
•Time
Social Context of Learning Environment
•Teacher-student relations
•Teacher-teacher-administration
•Peer group processes
•School-home
•Educational climate
STUDENT OUTCOMES
Academic Skills
Personal Development
Attachment to school and good behavior
STUDENT INPUTS
Individual student needs, abilities and interests
Student body context; school and classroom mix
Community factors

themselves facilitated by school organization and policy. Various combinations of school organization and policy will enable effective conditions to be established and support their continuation, while other combinations of school organization and policy will inhibit effective conditions from taking root and discourage their persistence. We represent this facilitating function in our model by a causal link between school organization or policy and classroom conditions. We have used the word "facilitate" rather than "cause", because the link between school organization and classroom conditions is an enabling or supporting connection rather than a determining connection. Effective classroom conditions, including excellent instructional practices and an appropriate social environment, can sometimes develop in many different kinds of schools using a variety of organizational structure and formal policies. For example, an outstanding teacher can usually function effectively in his or her own classroom regardless of the way the school is organized. Nevertheless, school organization and policy can facilitate the development of effective teaching.

There is practical as well as theoretical significance to the distinction we make between "school organization and policy" factors and "classroom operating condition" factors of effective learning environments. We believe it is possible to change either set of factors in a school improvement plan, but the changes in each involve different kinds of problems.

"Classroom operating condition" factors include instructional practices of teachers and social context variables such as interpersonal relations and normative climates. We have learned through recent efforts at school improvement that it is possible to directly improve the instructional practices of teachers through staff development programs with explicit training in improved practices (Gage, 1984). But because many of the improved techniques for teachers require professional judgments in constantly changing classroom situations (for example, discipline management techniques), the success in changing these operating conditions will depend in part on the professional capabilities of each individual teacher. Even more problematic are direct improvements in social context and interpersonal factors of classroom operating conditions. Workshop and staff development materials are available (Brookover et al, 1982) but their impacts on change have not been carefully evaluated, and even strong advocates of the "effective schools" movement will admit that social context factors such as "teacher expectations" or "the educational climate" are difficult to change directly.

Thus, one practical approach to school reform is to improve the proximate classroom operating conditions that are most important for student outcomes, but this approach must be able to directly alter teachers' (often idiosyncratic) instructional behaviors and the elusive informal and interpersonal context of instruction.

A second complementary practical approach to more effective educati is to work on "organization and policy" factors in the school. In our model, we list eight general factors in this category, ranging from school size and grade-span to policies on how to staff, schedule and group for instructional activities. While it may be no easier to change organizational and policy factors than to improve classroom operating conditions, the problems are different. With policy and organizational change, the problems are more likely to be political and bureaucratic (changing formal regulations, rearranging spheres of influence, overcoming inertia). With change of classroom practices, the problems are more likely to involve limitations of individual competencies or personalities.

But there are some clear advantages to making improvements at the level of policy and organization, because these changes often will last through the inevitable changes in personnel of a school, they often will facilitate improvements in a large number of proximate classroom conditions, and they often can be coordinated at the state, district and school levels. Making these improvements, though, requires a clearer understanding of how school organization and policy changes can facilitate improvements in the classroom operating conditions, and what practical approaches can induce reliable changes in the relevant school organization and policy factors.

The distinction between the school and classroom levels is, of course, more complex, because organization and policy factors can be at the classroom level as well as the school level (for example, instructional grouping practices in the classroom) and proximate operating conditions of learning environments can also involve school level factors (for example, relations between teams of teachers who come from different subject-matter departments).

We next discuss the specific variables within each of the major parts of our model, as well as the major potential causal relationships among the variables.

Student Inputs

We include student inputs in our model, not because these are alterable factors, but because the effectiveness of a school depends upon how well it is designed to meet the special needs, interests and abilities of its students.

We believe the design of effective schools requires close attention to student heterogeneity, to peer group influences, and to the socioeconomic composition of the student body.

Heterogeneity. For any given age, students will be at a wide variety of different stages of biological development, cognitive growth, and personal and social maturity. The extent of this heterogeneity grows greater with increasing age. A classroom may contain pairs of students of the same age and sex whose academic and extra-curricular performances less resemble each other less than they do others who are significantly younger or older than themselves.

This student diversity has important implications for the staffing and operation of schools. School organization and instructional policies may need to permit flexibility in learning arrangements and activities to facilitate teacher efforts to effectively engage the attention and effort of these diverse students.

The peer group. The peer group plays a powerful role in students' social and psychological development. The school may affect the direction and power of peer group influences (Epstein and Karweit, 1983) by the way it creates conditions for particular associations to form through the demography of classroom student assignments and extracurricular memberships. The strength of the norms of any single peer group can depend upon the number and variety of peer groups to which an individual is attached, which also may be affected by school practices that promote or allow student contacts. The way classroom rewards are structured, especially the interdependencies of student tasks and evaluations, may also affect the relative priority placed on academic and non-academic pursuits by student peer groups (Slavin, 1983).

Student-body composition. Student-body composition is the race, sex, and social class mix of the students enrolled in the school. Research has strongly suggested that student body composition is a major influence on the normative climate of a school (Coleman et al., 1966), and may constrain the types of policies and structures that can be established in a school (McPartland and McDill, 1982). Similarly, community influences can be important inputs that affect school programs, (Gottfredson and Gottfredson, 1985) including the level of support or opposition given by the community to school officials.

In research, it is always necessary to control for the influence of student background when estimating the effectiveness of school programs. But it is also important to attend to the ways that school programs and student inputs *interact* so that identical school programs produce dissimilar effects on students from different backgrounds. Thus, it is important to discover what *particular* learning experiences are most effective for individual students with *specific* needs or interests.

Student outcomes.

Several different classes of student outcomes are important for research to consider. Besides student learning of academic skills and attaining of academic competencies which is the main goal of education, student outcomes on a variety of measures of personal development and attachment to school take on special meaning as students get older.

Academic skills. Student learning of basic knowledge and skills in the main academic areas is a primary outcome. Instruction in academic subjects is the main purpose for which schools are established, and society expects the schools to accomplish this goal.

The curriculum requires attention to both basic skills in the mechanics of reading, writing and arithmetic and the development of higher order academic skills in the major subjects such as comprehension, problem solving, expository writing and critical thinking. There is some evidence of serious general problems of school effectiveness on both aspects of academic growth and development. For example, many local school officials believe their basic skills test score results show a noticeable drop-off in success at grade 5 and the subsequent middle grades that is not well-understood. At the same time, results from the National Assessment of Educational Progress on recent tests of mathematics, reading and writing suggest that improved performance may be especially evident on those exercised more easily learned by memorization and taught by rote drills, rather than on exercises calling for more complex thinking and measuring the ability to apply concepts to problem solving (National Center for Education Statistics, 1984).

Personal development. Schools may help students develop such psychosocial maturity outcomes as growth in self-discipline and self-regulation supported by a positive sense of self and well-developed internal goals, values, and standards.

In addition, one can think of a long diverse list of other non-academic talents and coping skills that are valuable to the individual and may be fostered in effective schools. These include abilities in music or the creative arts, leadership and interpersonal skills, work habits of industry and accomplishment, and coping skills in a variety of organizational and social settings. Our scientific knowledge is not well-developed of the numerous human non-academic talents that are needed and rewarded in adult life (e.g. Coleman, 1980). Consequently, researchers have not usually tried to measure non-academic talents in their studies of school effects (exceptions include the diverse checklists of student activities and accomplishments regularly used on National Merit and ACE surveys of college bound students).

Attachment to school and good behavior. How students react to their school life is an important outcome in its own right (Epstein, 1983), but it is also instrumental for other school effects on students. If students are chronically absent because they lack positive feelings toward the school, little learning is likely to occur. If a students are frequently having serious disciplinary problems with teachers and school officials, the school experience is not helping them develop the personal maturity needed later to successfully fill adult roles in work and community settings.

Chronic absenteeism by significant numbers of students occurs at the secondary school level, especially in large urban districts. Serious problems of violence and delinquency in school also begin in the middle-school years.

Drug and alcohol abuse, teenage pregnancy and the tragedy of teenage suicide are also serious problems where educational factors contribute to their later occurence. We have elaborated elsewhere how success or failure in school plays a unique role in the etiology of serious teenage problems (McPartland and McDill, 1977).

Classroom operating conditions.

The third of the four major elements of the model in Figure 1 is the group of "operating conditions" that directly affect student development, learning, and performance in classroom settings. We specify two broad categories of classroom operating conditions: specific instructional practices and the social context of learning.

Instructional practices include the way teachers design their lessons, deliver their instruction and manage their classrooms to create effective learning activities. A great deal has been learned in recent years on this topic as educational psychologists have established specific components of effective instructional practice in the elementary and middle grades.

The social context of learning includes the social and interpersonal conditions that operate during classroom learning activities. These conditions include student-teacher relations, peer group processes, relations among teachers and between teachers and administrators, school-home relations, and normative climates. Educational sociologists and social-psychologists have shown that these elements are important for effective schools, but much less is known about how to capture or direct these forces in schools than is known about how to train teachers in effective instructional practices.

Associated with Figure 1, we will discuss four different elements of instructional practices and five different elements of the social context of learning.

Instructional practices. Four elements of effective instructional design have been clearly identified: (1) quality of instruction (2) appropriate level of instruction (3) incentives for learning and (4) time utilization (Slavin, 1984; Rosenshine and Stevens 1984. Carroll, 1963; Karweit, 1982; Brophy, 1983; Anderson, et al, 1985; Brophy and Good (in press); Doyle (in press).

Quality of instruction is the degree to which the proper information or skills are presented to students in an appropriate form, sequence, and pace. Research evidence is clear that students learn more when the pace of instruction moves through more material in the same period of time, without sacrificing student comprehension. While content coverage is a strong predictor of achievement, we do not yet fully understand the interplay of classroom practices that maintain both a fast pace and a high rate of successful student mastery (Commission on Reading 1985, p. 88). However, many of the key elements appear to be potentially under the control of the classroom teacher -- such as establishing a clear plan of steps with appropriate materials for specific learning objectives, and the effective use of feedback fo increase student mastery.

The appropriate level of instruction is the degree to which material is presented to students at a level where individual students have the prerequisite skills to understand the material but have not already learned it. Some methods of targeting instruction go hand-in-hand with teacher practices that establish high quality instruction, such as the frequent assessment of students' level of mastery to link new material to previously learned concepts and to present materials that students can handle with a high success rate. But the appropriate level of instruction is also affected by the way students are grouped, whether whole-class, subgroup or individualized instruction. Again, researchers disagree about the efficacy of ability-grouping practices and the extent of the problems of management and motivation in using various individualized approaches.

Incentives for learning is the degree to which students are motivated to work on instructional tasks and to retain what is taught. Student motivation has been a major topic in educational psychology over the years and can be approached from a number of different directions (Ames and Ames, 1984). The sources of student motivation are many, but they usually include elements potentially under the control of the classroom teacher, especially in terms of how student performances are tied to formal and informal evaluations and rewards (Natriello and Dornbush, 1984). We now understand some useful motivational principles of incentive systems, especially related to the frequency and accessibility of valued rewards, and can incorporate these principles into teachers' classroom practice (Slavin, 1984). But, except for research on group incentives in the classroom and the use of home-based reinforcers, there has been little scientific study of alternative classroom evaluation and incentive systems as they affect student motivation and learning.

Time for learning is the degree to which students are given adequate time to learn what is taught. More time does not directly equal more learning, but more time in high quality instruction that is at the students' appropriate instructional levels will produce more learning. It is the "engaged time," when a student is productively involved in appropriate learning tasks, that matters. This depends upon the allocated time plus the teachers' skill in managing the class. Effective classroom management will minimize discipline problems and will minimize disruptions to learning activities from discipline problems that do arise. Skilled teachers will have efficient routines for managing potentially time-wasting chores that can accompany instructional activities (Karweit, 1983, 1984).

Social context of learning. Effective schools also include key elements of classroom processes that may not be so directly under the control of a classroom teacher as the components of instructional practice just described. These elements include, among others, the interpersonal relationships that occur among students and adults in the classroom, and the social climates that develop to produce different reputations and expectations for performance. The interpersonal aspects of learning environments have received special attention from many educators seeking to create schools that meet the developmental needs of their students (e.g. Lipsitz, 1985; Alexander and George, 1981), but few arrangements for changing interpersonal relations of students and teachers have been carefully evaluated.

For example, many critics of traditional junior-high schools focus on teacher-student relationships as a significant problem in these schools. The "middle school movement" developed in part as a reaction against formalized teacher-student relationships found in junior-high schools, which were seen as an outcome of modeling instruction for intermediate grade students on similar structures used in senior high schools. Instead of the typical secondary school's "subject-matter orientation" that emphasizes teacher expertise in a curriculum specialty, a "pupil-oriented" environment was called for to permit a closer teacher-student relationship to foster student personal development. The early adolescent is moving toward more self-regulation and autonomy during this period but, according to this view, still needs close personal contact with at least one adult in the school to support this growth. These more personalized and supervised teacher-student relations reflect typical elementary school practices, but advocates of this position expect teachers to also meet young adolescent needs for independence and self-direction. Little research exists on the dynamics of adult-student relationships that work well for the personal development of young adolescents who are at different stages of self-reliance.

Teacher-to-teacher relationships are also important aspects of the learning environment. The instructional flexibility needed to meet the diverse needs of students depends upon teachers cooperating with one another. Many educators advocate creating teams of teachers to work with shared student groups. They reason that teams can (1) more correctly diagnose individual needs and tailor learning experiences to meet those needs; and (2) more creatively develop lively learning activities that will appeal to children's and adolescents' sense of action, fun, and fantasy, which will better hold their attention and promote enthusiastic effort.

Advocates of teachers working as teams expect greater coordination of learning activities across formal subjects. However, no research has carefully and scientifically compared the actual advantages of different uses of teaching teams with the possible disadvantages that may occur if individual teachers do not get along, do not share common educational views, or do not effectively use common time to diagnose student needs and coordinate instructional responses.

Relationships among students in peer groups also can greatly influence the kind of learning environment that is created. Peer support is highly valued in the growing up process for most young people, but the peer influences can vary greatly in strength and direction for different individuals.

Which students an individual associates with may have a powerful effect on personal development. How many peer groups and close friends an individual is attached to may determine how influential any particular circle of friends may be. The overlaps among a student's associates in class, in extra-curricular activities, and outside of school may also influence peer group effects. Some research has examined these topics (Epstein and Karweit, 1983), but much more needs to be learned about how to coordinate the forces of the peer group to help students achieve academic and developmental goals.

The general climate or ethos of a school is also important to an effective learning environment. This factor includes the goals and normative expectations for behavior that develop in the school, and the collective reputatiion or image that can influence how an individual identifies with the school and is guided by its norms. Important dimensions of a school's climate include the strength and direction of the shared goals and expectations, and their clarity and consistency for subgroups of students and teachers.

Although there is agreement among researchers that more effective schools stand out from others in their educational climates, how different climates develop in schools that enroll similar student populations is not well understood.

School organization and policy.

To reliably create appropriate instructional practices and learning environments in schools, we need to understand how classroom operating conditions depend upon the enabling and support structures -- the organization of the school and its administrative policies. These school organizational and policy variables include school size, curriculum policies, staffing patterns and roles, grouping of students for instruction, scheduling, student moitoring and evaluation procedures, opportunities for student accomplishment, and grade-span.

To emphasize how schools may differ on these organizational and policy factors, it is helpful to compare the "typical" elementary and secondary school. Data should be collected on the actual distribution of school organization and classroom practice factors by educational level, since no reliable national data now exists on these matters. For this discussion, we will speculate on the differences between elementary and high school levels, and consider some other possibilities between these extremes. We will also discuss how each organizational and policy component may affect instructional practice or learning environments and what major research now exists on these effects. Table 1 was prepared to accompany these discussions.

School size: Size is a potentially important element of school structure at all levels because the number of students in a school can affect (a) student-teacher relations, (b) relations between teachers and school administrators, and (c) the types of peer contacts that result from classroom assignments and participation in extra-curricular activities.

Large schools may reduce the chances for positive student-teacher interactions, by making it less likely that students will feel closely supervised by teachers or that each student will develop a close personal relationship with an adult in the school (Garbarino, 1978; McPartland and McDill, 1977). Compared to a small school where most teachers would recognize most students by name, a student is more likely to "get lost" in the depersonalized environment of a large school.

Table 1

SCHOOL ORGANIZATION AND POLICY OF

TYPICAL ELEMENTARY SCHOOL AND TYPICAL HIGH SCHOOL

BEST COPY AVAILABLE

	SCHOOL ORGANIZATION OR POLICY	"TYPICAL" ELEMENTARY SCHOOL	"TYPICAL" HIGH SCHOOL
1.	Size	Small	Large
2.	Curriculum	basic skills, required courses courses	higher order skills, some electives
3.	Staffing assignments and roles		
	a. number of different students per teacher in a typical week.	one class (20-30 students) "self-contained classroom"	several classes (100 or more students)
	b. number of different subjects taught per teacher	all major subjects in the grades	one subject-matter specialty "departmentalized"
	c. number of different teachers per student	one	several
	d. Principal's role	instructional leader	school manager
	e. Advisory/guidance role	classroom teacher assumes responsibility for diagnosing student needs and providing or funding assistance	students are more responsible for seeking help when needed; adult guidance specialists are provided
4.	Grouping of students for instruction		
	a. homogeneous groups	within-class ability grouping	tracking and program differentiation (between-class grouping)
	b. low-achieving students	corrective instruction within class	separate remedial classes special teachers
5.	Scheduling		
	a. Time-schedule	flexible timing under control of teacher	fixed schedule of periods as students change classes
	b. student-schedule	intact classes of students remain together (self-contained classroom or block schedule)	students regroup for each period as they change classes
	c. required/elective courses, classwork	courses are required, teacher assigns classwork	combination of requirements and student electives of courses and classwork from teacher defined alternatives, occassional independent study
6.	Monitoring and evaluating students		
	a. grading practices	balance positive evaluation for both student performance and effort.	evaluations based on performance comparisons
	b. discipline	classroom variations in discipline management are permitted within school rules and procedures (personal authority)	school-wide rules, codes and procedures are in force (bureaucratic authority)
7.	Opportunities for student accomplishment	limited or no extra-curricular activities	extensive program of clubs, teams and activities

According to this view, students will more often be left to their own unsupervised activities in the anonymity of large schools.

Administrative practices are more likely in large schools to rely on bureaucratic processes, such as reliance on standard rules and regulations for governing teacher and student behavior. Relations between school staff members may be more impersonal and inflexible, which can create lower morale and an unwillingness by the staff to respond in innovative ways to problems (Garbarino, 1978). According to this view, in larger schools communication among staff is more difficult, school administration is more cumbersome, cooperation between faculty and adminstration in planning and implementing new programs is reduced, and clear well-understood policies are less likely. Some research does show smaller school size to be related to teachers' positive perceptions of school administration (Ebert, Kehoe and Stone, 1984; Gottfredson, 1985), to the absence of attendance problems among teachers (Winkler, 1980) and to cooperative educational activities among the teaching staff (Bridges and Hallinan, 1978; Bridges and Hallinan, 1978).

Smaller schools may also produce different peer groupings of students because extra-curricular offerings and tracking practices are affected by the size of the student body. Research at the high school level has shown that small schools often induce a higher percentage of students to get involved in extracurricular activities (Baird, 1969; Barker and Gump, 1964; Grabe, 1981; Kleinert, 1969; Wicker, 1969), which in turn connects the average student to a more diverse personal network of peers.

The track levels in a small school may also expose each student to a more diverse set of peers than would occur in a large school with the same range of abilities in the student body. A large school that tracks students according to test scores or previous academic achievement will often create more homogeneous classes with greater differences between the top and bottom classes in student abilities than will be found in a small school with the same range of student abilities. This occurs because the small school have fewer sections of each course, so each class will be closer to the mix of students in the school at large. Also, there will be fewer very bright or very slow students in a smaller school to be assigned to exclusive classes. Thus, when the student mix in the school is about the same in large and small schools, extracurricular activities and tracking will often produce more diverse peer contacts in the smaller schools.

On the other hand, larger schools will ordinarily draw from a larger and more diverse attendance area, so the student body as a whole is likely to be more heterogeneous by family background, race and other characteristics. Whether an individual student in a large school will actually come into contact with peers from different backgrounds depends in part on school policies concerning tracking and extra-curricular activities. More direct research is needed on how school size changes the opportunities for diverse peer contacts and the actual formation of peer groups.

Large schools may have some advantages in variety and quality of curriculum offerings (Conant, 1969). A large school is more likely to have enough students interested in certain unusal courses -- such as some foreign languages, technical courses, or advanced level offerings -- to justify the expense of staff and equipment to provide these courses.

Some structures may offset the disadvantages of the impersonality of large schools by creating smaller identifiable units within a section of the same building: a "school-within-a-school." In these subgroups, teacher and student assignments and activities emphasize identification with the smaller unit and its members. Educators have suggested various ideas about how to conduct activities with and between different subunits of a large school to achieve good interpersonal relations and student attachment. Little careful research has been conducted on these issues.

Staffing patterns and roles: Elementary school teachers are usually assigned to a self-contained classroom where a single teacher is in charge of instruction in all major subjects for an intact class of students. High school teachers usually are "departmentalized" by subject-matter area, and assigned to teach courses in one specialty area of the curriculum to different classes of students during the school day and school week. Teacher certification regulations in most states reflect this difference: Teachers are certified by level at the elementary grades without subject-area distinction while teachers are certified by subject-matter specialties at the high school grades.

Thus the typical elementary school student receives almost all instruction from one home-room teacher, while the typical high school student receives instruction from several different teachers. And each elementary teacher is responsible for a single class of 20 to 30 students but must prepare lessons in a number of subjects, while each high school teacher may see well over 100 students in a given week but can concentrate on teaching in one curriculum specialty.

Both the high school and elementary modes of staffing offer advantages and disadvantages. Departmentalization and subject-matter orientations may increase the quality of instruction by allowing teachers to increase their competence in a curriculum speciality, provide outstanding learning activities for a limited number of separate daily preparations, and bring a special enthusiasm to particular areas of the curriculum that is sustained by departmental colleagues. On the other hand, the self-contained classroom of the elementary grades is believed to achieve strong "pupil-orientation" due to the close and concentrated associations between a single teacher and a fixed small group of students.

Various structures have been proposed to strike a balance between these two poles -- to achieve a personalized learning environment while allowing individual teachers to develop high quality curriculum specialties. These include a variety of teacher team arrangements, such as special roles for the homeroom advisor in a modified-departmentalized school. (See, for example, Alexander and George, 1981, Chapters 4 and 5.) For example, a team of two or three teachers could serve 50 to 75 students, with each teacher specializing in a broad curriculum area such as math and science or language arts and social studies. The team would work together to diagnose student needs, establish student instructional groups, and coordinate and schedule learning activities. Provision for team planning time would be built into the weekly schedule. Another example would be to have larger teams of subject-matter specialists (5 or 6) serve larger common groups of students, but also provide a personalized guidance program in which each student is assigned to one particular teacher in an advisor-advisee capacity.

Although creating teacher teams with scheduled time to diagnose students and plan instruction would seem to be a way to achieve desirable instructional practices and interpersonal environments, there is no guarantee these opportunities will be taken advantage of by teachers. How much does the success of teacher teams depend upon how well team members like each another, on how the use of team planning time is supervised, on the roles defined within the team, or other operational factors? Research has not carefully investigated these questions, or other issues related to staffing patterns and the effectiveness of the resulting learning environments and instructional practices. We do not even have good descriptive statistics on the use by schools of different staffing patterns at the elementary, middle and secondary levels.

The role of the principal is another key issue of staff roles where interesting comparisons have been made between elementary and high schools (Farrar, Neufeld & Miles, 1984; Firestone & Herriott, 1982; Purkey & Smith, 1985). The elementary principal is more frequently seen primarily as an instructional leader (assisting and evaluating individual classroom teachers; establishing the school climate). The high school principal is more frequently seen primarily in an adminsistrative or bureaucratic role, who helps maintain priorities on learning, is in charge of the rules and their enforcement, and involves teachers appropriately in decisions on school-wide matters.

Grouping students for instruction: Elementary schools are more likely to randomly assign students to classes but to group students within the classroom for instruction. Within-class ability grouping usually involves creating three homogeneous groups of students who have similar current levels of achievement. Within-class ability grouping in elementary school rooms is almost always used in reading instruction and is often used in math instruction. (Peterson, Wilkinson and Hallinan, 1984).

In contrast to elementary schools, high schools create more homogeneous instructional groups by placing students in programs and tracks by achievement level. High schools separate students according to entire programs -- such as academic or college preparatory, general, vocational or technical, and commercial or business -- and according to course track level within the program -- such as high, medium, and low sections of English courses at the same grade. Because students are placed into separate programs and courses by achievement levels, there is rarely an use of within-class ability grouping at the high school level.

The approach to corrective instruction for low-achieving students in elementary and high schools usually mirrors the above grouping patterns. Elementary schools frequently use with-in-class corrective instruction; high schools separate remedial classes with specialized teachers. However, elementary schools have often used Federal support (Title I and Chapter I funds) to provide separate remedial teachers and programs that pull out low achievers from their regular classrooms for special instruction.

Judging from the volume of published articles, tracking and ability grouping would appear to be among the most thoroughly researched topics in education. And there seems to be considerable agreement among researchers about the impact of grouping practices: a poll of researchers would probably show that a large proportion believe that grouping practices have been proved to have a negative effect on the development of students in the lowest groups. <1>. But, because there have not been many randomized experiments in research on tracking or ability grouping, and because correlational research cannot convincingly control for the different student rates of learning that are usually related to the group assignments, we believe it is too soon to draw scientific conclusions about the effects of alternative grouping practices. Additional experimental research is needed to compare different methods of grouping students for instruction, and this research must pay special attention to how actual classroom practices are adapted to different instrutional groupings.

The research should examine detailed practices within various subgroups and establish convincing scientific controls on initial student differences. Recent studies of elementary and junior high classes by Johns Hopkins researchers and others strongly suggest that certain grouping practices, accompanied by appropriate classroom activities, can yield learning benefits for all levels. (Slavin and Karweit, 1984, 1985; Doyle, 1984; Evertson, 1982; Filby et al, 1982; Bossert et al, 1984)

While there is much consensus about grouping practices at elementary and high schools, good data on grouping practices at the middle grades is not now available, even to describe the distribution of alternative approaches in our nation's middle and junior high schools. We need surveys to determine how middle grade classrooms now use or don't use between-class tracking and/or within-class ability grouping. We also need to address major questions of relationships and causality. How do

---

<1> We also suspect there is majority agreement among teacher about the effects of grouping, but with an opposite conclusion. We predict that a majority of teachers would report their belief that homogeneous grouping of student produces greater learning because instruction is targeted to students' current needs.

middle-school grouping practices depend upon the subject-matter area of the course and the staffing patterns in use? How do intermediate grade teachers adapt their instructional practices to the needs of each group, including provision for student initiative in learning tasks? What relationships between students and teachers and among students develop with different instructional grouping arrangements? What student outcomes are more likely to occur when alternative grouping structures are coupled with particular classroom practices?

The opportunities created by alternative instructional groupings for more efficient instruction depend upon the actual tasks established in the classroom (Bossert, Barnett and Filby, 1984; Bossert and Barnett, 1981; Bossert, 1979). Bossert's analyses of how within-class ability grouping may be associated with a variety of actual classroom practices can be extended to questions of how between-class grouping level may foster different instructional management and learning environments within the classroom. The goal of this research is to show how the positive effects of particular grouping policies depend upon the actual classroom practices that are used to take advantage of the policy.

Scheduling: The manner in which school schedules are made, --dividing the school day into different periods for instruction in separate courses and assigning teachers and students to different classroom locations during the day -- is closely related to the decisions on staffing and instructional grouping we have discussed.

In elementary schools, the teacher in a self-contained classroom usually schedules the mix of time devoted to each curriculum area, within general guidelines provided by the school. There are no fixed periods announced by the ringing of bells throughout the day to signal the time for teachers of students to change locations. Teachers use their professional judgment to arrange instructional activities of different durations to fit the changing demands of curriculum topics or changing needs of students.

At the high school level, the forms of tracking and staffing used will be related to the type of scheduling. Since teachers must meet different classes throughout the day and since each student receives instruction each day from different teachers, a centralized schedule is necessary to divide the school day into fixed periods of time and provide lists to locate where each teacher and each student is to be each period for instruction in specific courses. The change in periods is usually announced by the ringing of bells and signals the movement of teachers and students through the school corridors. These regular sounds and traffic have come to symbolize for some the over-regimentation of the American high school.

The schedule may vary in complexity depending upon the flexibility of instructional time periods, the variability of student groups, and the number of elective courses permitted, as well as other local complicating circumstances. For example, a "block schedule" keeps student groups together as intact classes for most periods of the day with departmentalized staffing providing different specialists to teach each course. This approach is used in high schools that assign students to programs and tracks based on a single test score or criteria. Other high schools may design more complex student schedules that allow each student to be grouped differently in each major subject, or to permit elective courses for students at several points in the week. Other high schools may define instructional periods by smaller modular time units so some courses can be formed using multiple modules to provide more instructional time on certain days.

Appropriate scheduling for the lower grades depends upon decisions that are made about staffing, grouping and curriculum flexibility. Some middle school educators advocate using (a) interdisciplinary teaching teams to achieve curriculum quality and flexibility, with (b) student assignments that keep classroom-sized groups together as a unit for most of the day to establish a more secure peer group identification for each student. (See, for example, Alexander and George, 1981). One suggested scheduling structure is a modular version of the blocked schedule that allows time for teacher team planning and coordination. Other scheduling structures exist to help accomodate particular goals of curriculum and learning environments. (See, for example, four interesting cases of middle school operations described in Lipsitz, 1984.)

We agree with educators that scheduling structures should be devised to best support the desired instructional program and learning environments, and should ordinarily follow decisions about the approaches to curriculum design, staffing and instructional grouping that are expected to produce these practices and environments. But designing and applying various structures of staffing, grouping and scheduling, and testing their impact on student outcomes are matters for direct study and scientific evaluation.

Student monitoring and evaluation procedures: The "pupil-orientation" of the elementary school and the "subject-matter orientation" of the high school may also be reflected in procedures for monitoring and evaluating students through policies on grading, discipline, and advising. Just as we expect many elementary teachers in self-contained classrooms to have more latitude for scheduling time for different instructional topics and more flexibility in grouping students for alternative learning activities, compared to departmentalized high school teachers following centralized schedules, we expect elementary teachers to be less affected by school-wide constraints on grading, disciplinary and advising practices. Since elementary teachers are thought to be more "pupil-oriented," we expect their grading practices to strike a more conscious balance between a student's effort at school work and a student's rank-in-class on tests and other measures of academic performance, so that a low achieving student who tries hard would receive some positive feedback. Grading at the high school level is more likely to be looked at as a way of sorting students that as a way of motivating students. Grading in secondary schools is also often tied up with tracking practices, where different floors and ceilings for assigning marks are maintained in low or high track classes.

Opportunities for student accomplishment: High schools provide a wider range of ways that students can gain recognition than do lower level schools. Extra-curricular activities in high schools permit many individuals to develop and demonstrate competence in athletics, in musical and artistic performance, in producing a school newspaper or yearbook on some publication outlet for creative writing, in working with others on projects that require organizational and interpersonal skills, in assuming decision-making or representative roles where political skills are useful, and in enjoying a variety of special interest and hobby activities where individuals can develop unique knowledge or skills. Various reports on high schools have called for expanding the opportunities for students to assume initiative and responsibility by providing services to others or their community (Boyer, 1983; Coleman, 1974; Newmann, 1981).

The need for a wide range of opportunities for accomplishment is also vital for middle-school students, because early adolescence is a time of striving for achievement and competence. To develop a positive self-concept and to mature in self-confidence, young adolescents need to develop general abilities to function successfully in a variety of situations and expand the particular talents that bring them special pleasure or accomplishment.

To identify ways to provide more opportunities for student accomplishment, research is needed on both (a) how to make rewards for *academic* development accessible to more students, and (b) how to provide a *wider range* of activities that require and recognize a diversity of human talents. For the first question, we need to study alternative academic evaluation systems that are responsive to individual effort, improvements in performance, or alternative modes of demonstrating competence. For the second question, we need to identify a wide range of extra-curricular, co-curricular and service activities and evaluate their effects on student development and self image to provide a knowledge base for expanding the range of opportunities for student accomplishment in schools.

Grade-span: There is also a school organizational question that must be met on the district level rather than approached within each school: the appropriate structure of grade-span organizations for schooling children between ages 5 and 17.

For example, there has probably been more written on the advantages or disadvantages of different grade-span structures for the intermediate grades with less clear research guidance than any other single topic about schools for young adolescents. In 1983, the Educational Research Service (ERS) published a 200-page summary of research on the organization of the middle grades that used 424 separate references. Few consistent findings emerged from the review. The ERS conclusion that "the quality of the school program is more important than grade level organizations" echoes our view that structural features such as grade-span are important only in so far as they enable or support those instructional practices and learning environments that produce desirable student outcomes.

We believe further research is warranted on how alternative grade-span structures may be related to instructional practice, learning environments and important student outcomes.

First, past research on these issues has usually not applied careful statistical controls on differences in student inputs. Second, there is a reasonable basis to think that grade-span may directly affect some variables, especially peer group environments and their consequences. For example, advocates of the 6-8 grade-span argue that ninth graders are much closer developmentally to tenth graders than to eighth graders, and a less desirable ("too sophisticated") environment is produced when ninth graders are left in the same school with younger adolescents and pre-adolescents. More careful research is needed on peer group reference groups and educational climates in schools with different grade-spans (Blyth, Hill and Smyth, 1981).

Third, a small number of well-designed studies find impressive effects of schools with different grade spans on good school attendance (Slavin and Karweit, 1982) and self-esteem (Blyth, Simmons and Bush, 1978).

Fourth, other grade-span sequences may deserve careful examination for their support of learning environments that are well suited to the needs of young adolescents. The traditional two-stage 1-8 elementary and 9-12 secondary school structure should be carefully studied -- it remains the modal pattern today among private schools.

II. A Design for Research

Some General Design Considerations

The model presented in the previous section is a comprehensive account of instructional and organizational choices made by schools and school districts that are likely to have important consequences for student academic achievement, personal growth, and school-related attitudes and social behavior. As such, a single research project or research design could never adequately measure and test all of the parameters and hypotheses in the model.

A mixture of research designs and projects is clearly needed. An appropriate mixture would include systematic observational studies, so that "dense" measurements can be made where quick-and-dirty survey questions would produce severe distortion. It would include repeated measurements on the same students over many years, so that longer term outcomes of schooling could be followed. And it would include using multiple instruments and multiple respondents at the same site, so that special expertise and complementary perspectives would contribute to the final information product.

Yet it would also be useful for the multiple projects and designs to build upon one another, using the same sample of schools and students, and the same theoretical framework, so that each project could gather data not only to answer questions that it posed but so that it could inform remaining projects in the series, avoid duplicative data gathering, and optimise the data collection method to the kinds of information needed.

Particular research questions drawn from the common theoretical framework may require specific variations in a common study design. However, following the model's emphasis on school and classroom factors as independent or "treatment" variables and student academic performance, personal development, and school-related attitudes and behavior as outcome variables, we see the following as general considerations that should be followed for much of the research that is needed.

First, the analytic unit to most of the specific studies should be the *school* or the *classroom,* not the individual student. Where the concern is with school-level policies that affect instruction and learning, such as between-class ability-grouping, or departmentalization of the teaching staff, the unit of analysis is the school. Where classroom instructional practices are the concern, the unit of analysis is the individual classroom. Even where long-term outcomes are the issue, the students followed over a several year period continue to be valid measurement points for the study of the consequences of *school organizational* or *classroom instructional* treatments. Longitudinal student data is merely the means by which appropriate data is gathered to study the effectiveness of school and classroom practices and conditions.

Secondly, it is also appropriate that the sampling unit be the school and sub-samples of its classroom groupings (e.g., 5th period, room 120, Monday). Again, the student-based survey instruments are a means of obtaining measures of the effect of the classroom treatment, whether this be the grading and incentive practices of the teacher or practices of school-home cooperation.

Third, measurements of school policy and classroom practice need to be made at appropriate points and with sufficient investment so that measurement error on individual cases is reduced to a manageable point. Asking teachers or school administrators to describe policies will be generally accurate if those policies are conscious, public, stable, and explicit. The more that questions deal with behavior patterns determined by custom, internal politics, or general agreement, and the more that the patterns vary according to the characteristics of the specific instance, the more that attention has to be paid to obtaining multiple sources of data about the factor in question, measuring the behavior or policy at different points during the school year, and using judgments of external observers rather than relying solely on self-reports of school practitioners.

One of the most important "givens" of school research is that schools and classrooms providing differing treatments often start with student groups that hardly resemble one another. Thus, it is important to devote resources to measuring both the "starting" and the "ending" attributes of students that may be affected by the type of school or classroom treatment being provided. Although simple cross-sectional survey designs have been the norm for most studies of school "effects," our understanding of the consequences of differential educational treatments has been immeasurably harmed by using "socio-economic-status" and other background variables to stand in as surrogates for student "starting" characteristics. It is time that all major studies of school practices be done with both "pre-test" and "post-test" points of measuring student outcomes, at the very least.

Not only do students have differing backgrounds that affect school performance and attitude, but each student has a prior history of school experiences and exposure to various school and classroom practices. A study beginning at one point in time should also, if possible, attend to issues of prior school experience, including gathering data about previous teachers and previous schools attended.

In analyzing the effects of instructional practices in individual classrooms, research must take into account that in most secondary schools and in many elementary schools, students are taught in more than one classroom setting by multiple teachers applying different practices in different ways. Studies of the impact of classroom practices must consider how these simultaneous multiple treatments are likely to affect the outcome measures of interest. Often, it may be valuable to obtain comparable survey instruments about classroom practices from each teacher in the school who also teaches some of the students who are in the "sampled" teacher's classroom.

Many questions about the consequences of schooling relate to student adjustment, attitudes, and behavior in subsequent school and work experiences--for example, questions about how experience in one school affects students' disciplinary habits, school attendance, and school performance at a subsequent school attended. Such questions require that the students be followed for at least several years beyond the treatment being studied. In addition, for many schooling processes, it may be that only consistent treatment applied in successive schooling environments has important and measurable consequences. Thus, where such "small, cumulative impacts" are hypothesized, it is important that the research design plan from the start to be a longitudinal one, wherein similar school and classroom treatment variables are measured on successive occasions along with student outcome variables.

Finally, we need to mention one other important consideration that should guide the design of a research plan for studying the effects of school and classroom practices on student outcomes. Although the most important questions on this topic are causal in nature--how do different organizational and instructional practices affect the achievements, attitudes, and behaviors of different groups of students--we still lack basic descriptive data about the factors discussed in the model.

An important contribution of N.C.E.S. would be to provide this descriptive information as part of an effort to understand how these factors affect schooling outcomes. We need descriptive information about how instruction is organized and conducted for different subjects and at different grade levels in different kinds of schools serving different student populations. The need for descriptive data requires that increased attention be given to sample representativeness and sampling strategy. <2>

A Suggested Research Design.

The theoretical model guiding this discussion is a broad one, implying a wide variety of research needs and plausible research strategies, and covering the full range of organized schooling from kindergarten through college. Even the design for an "umbrella" study, under which specific research questions on specific school and student populations could be studied in more detail, must leave out some of the possible topics and coverage of schools and students. What we suggest below, then, constitutes a selection from among the universe of designs that might inform the questions posed in our model.

Because we are interested in school and classroom "treatments," we propose sampling schools and classrooms rather than students. However, because our interest is in outcomes for groups of students of particular ages or grade levels, stratified samples need to be drawn that take into account student grade levels, and student data should be collected for students in a particular grade in sampled classrooms.

Five major goals affect the choice of schools, classrooms, and students to be sampled. The first is to maximize the variety of types of schools (grade-level ranges, student size, public vs. non-public control) included in the sample. This goal suggests that stratifying schools by size, grade-span, and control, and drawing samples of similar size for each stratum would be preferable to solely sampling schools with probabilities in proportion to their size.

The second goal is to obtain descriptive data on the methods of classroom instruction used at as many grade levels as practical. This goal suggests that at least teacher data be collected from teachers of all grade levels.

A third goal is to measure the impact of school-level treatments on students. This suggests sampling classrooms with students in their first year of being exposed to such a treatment--e.g., sampling 7th grade students in 7-9 junior highs, and 6th grade students in 6-8 middle

---

<2> By itself, causal studies are less dependent upon having a representative sample because associations between school practices and outcomes and causal relationships involving these variables are likely to be much more stable over different sub-populations (of schools) than are descriptions of school characteristics and practices themselves.

schools. This maximizes how long each student sampled will be "exposed" to the school's treatment during the study period.

A fourth goal is to measure the effects of classroom instructional practices on students. In the many schools in which students are taught by more than one teacher, this requires that teacher data be collected from each of the teachers whose instructional patterns might affect individual students whose outcomes are being measured.

A fifth goal is to obtain measures of school effect on student outcomes measured across transitions to new schools and schooling levels. This goal suggests sampling students so that as many as possible will have moved to a new school or non-school environment for a similar, limited duration (e.g., one year) prior to a follow-up survey. If, for example, we sample students in their first year at a school with three grades, we could follow-up these students, say, in base-year plus 4--that is, in the second year after entering their next level of schooling.

Although some of the goals suggest somewhat contradictory principles for choosing measurement points of classrooms and students, the following design maximizes attainment of the five goals as much as possible. In particular, it takes into account the number and size of schools of various grade-level ranges to maximize diversity of school selection; it maximizes the length and purity of a school "treatment"; and it maximizes the number of students for whom we can measure the impact of school and classroom treatments on a school transition experience.

According to this design, schools would be stratified into the following groups according to grade-span and student enrollment per grade-level. (See Table 2.) In addition, the sample would be further stratified by public vs. non-public control, although this is not reflected in Table 2 below. Sample sizes for these strata need not be identical--that is, other factors may need to be considered as well--but the sizes should reflect a primary interest in obtaining as diverse a sample as possible along the stratification dimensions.

School-level data should be collected from appropriate administrative persons at each sampled school. Teacher data--both self-reports of classroom practices and "informant" data about school conditions--should be collected from simple random samples of the full-time teaching staff. The sample sizes per school should reflect both a minimum number (e.g. 10) and decreasing fractions for larger schools. In addition, however, the sample of teacher data should be supplemented in order to obtain reports from each teacher responsible for the instruction of students in the classroom(s) sampled for longitudinal follow-ups (see below).

Classrooms selected for studying the impact of school organization and classroom practices on student outcomes should be selected accordi to the grade-level of the plurality of their students; they should be

Table 2:

Number of Schools Serving Grades 5 - 12, by grade span, by enrollment per grade*

Grade Span	Number and Percent of Schools by Enrollment Per Grade				
	<25 students per grade	25 - 70 per grade	71 - 150 per grade	151+ per grade	Total Schools
PK,K,1 - 12:	3592 (64%)	1779 (32%)	197 (4%)	12 (0%)	5580
PK,K,1 - 8:	7311 (47%)	7243 (46%)	994 (6%)	32 (0%)	15580
PK,K,1 - 6,7:	3911 (16%)	13981 (56%)	6612 (27%)	254 (1%)	24758
PK,K,1 - 5:	738 (7%)	5586 (51%)	4395 (40%)	227 (2%)	10946
4 - 6:	20 (2%)	155 (16%)	510 (53%)	273 (28%)	958
5 - 8:	48 (4%)	260 (22%)	588 (50%)	281 (24%)	1177
6 - 8:	26 (1%)	286 (7%)	1072 (26%)	2774 (67%)	4152
7 - 8:	13 (1%)	126 (5%)	389 (14%)	2238 (81%)	2766
7 - 9:	12 (1%)	74 (3%)	281 (11%)	2148 (85%)	2515
7 - 12:	580 (17%)	1503 (45%)	986 (30%)	275 (8%)	3344
8 - 12:	70 (13%)	130 (24%)	187 (35%)	154 (28%)	541
9 - 12:	644 (6%)	1758 (17%)	2598 (25%)	5535 (52%)	10535
10 - 12:	45 (2%)	151 (7%)	299 (14%)	1723 (78%)	2218
Other spans incl. 5 - 12:	1086 (21%)	987 (19%)	1317 (26%)	1747 (34%)	5137
Spans incl. PK - 4 only:	1002 (12%)	2793 (34%)	3276 (40%)	1017 (13%)	8088
Total, U.S.**	19098 (19%)	36812 (37%)	23701 (24%)	18690 (19%)	98301

---

* Source: Data tape of U.S. School Universe, 1984, Quality Education Data, Denver, Colorado.

** Excludes 3,348 schools classified as voc-tech, alternative, or special education and others for which grade spans were not available.

restricted to classes of academic subjects; and they should be sampled in proportion to the number of hours during the school year that the students meet that class. At each school, only specific grade-levels should be sampled; and these should be determined by the grade-span at the school. Our preference for the grade-levels to be sampled is shown in Table 3. This choice takes into account both the goal of sampling students early in their experience at the school and the goal of providing for follow-up studies to be conducted at appropriate points in the students' schooling careers (see below).

Table 3:

Grade-Levels Sampled by Grade Span of the School

Grade-span of school	Grade-level of classes sampled (Base year)	Grade-level in Year 3 follow-up (Base yr. +2)	Grade-level in Year 5 follow-up (Base yr. +4)
P,K,1-12	4,7,10	6,9,12	8,11,(14)*
P,K,1-8	4,6	6,8	8,(10)
P,K,1-6,7	4	6	(8)
P,K,1-5	3	5	(7)
4 - 6	4	6	(8)
5 - 8	6	8	(10)
6 - 8	6	8	(10)
7 - 8	7	(9)	(11)
7 - 9	7	9	(11)
7 - 12	7,10	9,12	11,(14)
8 - 12	8,10	10,12	12,(14)
9 - 12	10	12	(14)
10 - 12	10	12	(14)

* Parentheses indicate grade levels that are beyond the grade level range for the base year school.

We acknowledge that sampling classrooms (and thus collecting longitudinal data on students) according to the grade-span of the school is a design that conflicts with the pattern used in previous studies. However, we strongly feel that the research design should be determined by the substantive questions; that the important substantive questions concern the impact of school and classroom practices on students; and that this emphasis requires that the school grade-span rather than some arbitrary choice of grade levels should determine the selection of student groups whose outcomes are to be measured over time.

In addition, in order to measure the diversity of instructional practice, the notion of a "classroom" should be defined so as to incorporate a cluster of teachers who provide "teamed" teaching--that is, a coordinated teaching practice--to a common group of students. Classrooms (or clusters of classrooms) should be sampled inversely in proportion to the number of teachers involved. Thus, a team of five teachers teaching a group of 125 students would be sampled together, but with only 1/5 the probability of an individual teacher teaching a self-contained group of 25 students.

During the base year, each classroom should be studied near the beginning of the school year, and again near the end of that school year, with teacher questionnaires, student questionnaires, and possibly a classroom observation instrument. Then, each student who was a member of a sampled classroom at both points during the first year would be followed up on two occasions--once near the end of the second school year following the base year and once near the end of the fourth year after the base year. Again, the primary instrument at these followup points would be student questionnaires, but teacher practice and school organization survey instruments could also be employed, funds permitting.

Using the initial selection of grade-levels according to school grade-span proposed in Table 3 above, the third year followup (base year + 2) will occur for most students at the end of their final year at their base year school. The fifth year followup (base year + 4) will occur for most students in the second year after their transition to a new level of schooling.

III. Selected Questionnaire Items

The variables that should be measured in these surveys are those listed earlier in our discussion of the model in Figure 1. For school structure and classroom processes, the survey should include indicators of staffing patterns, grouping practices, scheduling, monitoring and evaluating students, opportunities for student accomplishment, as well as teacher-student relations, teacher-teacher relations, and educational climates. In addition, the surveys should contain measures of peer group processes and normative environments, and student outcome measures of academic skills, personal development, attachment to school and good behavior.

It would also be valuable for the survey to obtain attendance and discipline records and academic achievement score results for each student sampled. Because of the technical problems of establishing comparability between different achievement tests used at different times in various schools systems, it would be best to use a short (20-30 minute) achievement test using NAEP items administered to students as part of a 45 to 60 minute survey. (See Messick, Beaton and Lord, 1983, p.79 on use of NAEP items). Various alternatives could be considered, including basing sampling decisions on the type of test available in school files.

With a special sub-sample of schools, the survey should be accompanied by two-day site visits to draw a narrative profile of the school and its operation. Our model for this activity is the recent book by Lipsitz (1984) that includes detailed narrative descriptions and analyses of four interesting schools for young adolescents.

The following are some questions that might be used in a study of the impact of school organization and classroom instruction on student outcomes. The questions included here represent only a limited portion of the survey items needed. They are aimed primarily at measuring between-classroom and within-classroom grouping practices, scheduling of students and teachers, and arrangements for teaming or clustering of instructional groups.

Selected Questionnaire Items for Principals of a Middle-School

(Questionnaire items would differ to some degree according to the range of grade levels at the school.)

---

1. Do students at your school stay with the same class group for all academic subjects (English, Math, Social Studies, Science), or do they attend different classes with different groups of other students? (CIRCLE ONE CODE FOR EACH GRADE LEVEL AT YOUR SCHOOL.)

	Students Stay With The Same Class Group For...		
	All Academic Subjects	Some Subjects But Not Others	None (Each Subj., Different Groups)
Grade 6	ALL	SOME	NONE
Grade 7	ALL	SOME	NONE
Grade 8	ALL	SOME	NONE

---

2. For which acadmic subjects are most students assigned to classes by ability (so that some classes are higher in ability than others) FOR EACH GRADE LEVEL, CHECK ALL SUBJECTS GENERALLY ORGANIZED BY ABILITY.

For Which Subjects Are Classes Organized by Ability?

CIRCLE ALL THAT APPLY.

Grade 6:	ENGLISH	MATH	SOCIAL	SCIENCE STUDIES	NONE
Grade 7:	ENGLISH	MATH	SOCIAL	SCIENCE STUDIES	NONE
Grade 8:	ENGLISH	MATH	SOCIAL	SCIENCE STUDIES	NONE

---

3. Some schools organize academic classes to be composed of students of more than one grade level. Other schools place students of each grade level in their own classes. Please answer about each combination of grade levels at your school? CIRCLE ONE CODE PER LINE.

How Many of the classes attended by ...	...also have at least several...				
6th graders	7th graders	MOST	MANY	FEW	NONE
6th graders	8th graders	MOST	MANY	FEW	NONE
7th graders	8th graders	MOST	MANY	FEW	NONE

---

4. At your school, what kinds of teaching assignments do most teachers of the major academic subjects have? CIRCLE ONE CODE FOR EACH GRADE LEVEL.

	Self-contained: each teacher teaches all subjects to the same students	Departmentalized: each teacher teaches the same subject to several different classes of students	Mixed: teachers teach some but not all subjects to the same students
Grade 6	SELF-CONTAINED	DEPARTMENTALIZED	MIXED
Grade 7	SELF-CONTAINED	DEPARTMENTALIZED	MIXED
Grade 8	SELF-CONTAINED	DEPARTMENTALIZED	MIXED

---

5. Some schools use team scheduling in which, for example, four teachers of different subjects teach the same four classes of students. Does your school use this scheduling method for students? (ANSWER FOR EACH GRADE LEVEL.)

	Team Scheduling Used?	
Grade 6	YES	NO
Grade 7	YES	NO
Grade 8	YES	NO

IF "YES" FOR ANY GRADE LEVEL:

5a. Is there a specific planning period set aside for each group of teachers who work together?

YES NO

Selected Questionnaire Items for Teachers of a Middle-School Class Sampled in the Survey

---

1. Which of the following best describes your current teaching assignment? (CIRCLE THE NUMBER TO THE RIGHT OF YOUR CHOICE.)

I teach one group of students for the entire day..............	1
I teach one subject to several different classes of students..	2
I teach several subjects to more than one class of students...	3

---

2. In the table below, list the subjects, student grade-levels, and general ability levels of the classes that you teach during the week. List only academic classes such as English, Math, Science and Social Studies, and their specialties. LIST EACH SUBJECT ON A SEPARATE LINE, even if you teach them to the same class. AND LIST EACH CLASS ON A SEPARATE LINE, even if you teach the same subject to different classes of students in the same grade.

SUBJECT OF CLASS EACH SUBJECT & CLASS ON A SEPARATE LINE	HOURS PER WEEK	STUDENT GRADE LEVELS (K-12)	ABILITY LEVELS High=Hi, Average=Av, Low=Lo, Mixed=Mx (CIRCLE ONE PER LINE)
a)			Hi Av Lo Mx
b)			Hi Av Lo Mx
c)			Hi Av Lo Mx

etc.

---

3. The remaining questions in this survey concern only one of your classes. The class is shown on the cover of the booklet. by the day-of-the-week and time-of-day that you meet it. Which entry in the above table is for the class and subject that you teach at that particular time? (WRITE ITS LETTER -- "a", "b", etc.)

LETTER (Q.2) OF THE CLASS SAMPLED FOR STUDY: _________

---

4. Circle all of the other subjects that you teach to the same group of students who are in the sampled class.

English...........................	1
Reading...........................	2
Mathematics.......................	3
Science...........................	4
Social Studies....................	5
None Other: only one subject to this class.....................	6
Other (specify):_______________	7

---

5. Sometimes teachers divide their class of students into groups for instruction based on their demonstrated abilities. Do you do that for any subjects which you teach to the sampled class? (CIRCLE "YES" OR "NO".) If "yes," circle the subjects for which you ability-group for this class.

NO: DO NOT ABILITY-GROUP FOR THIS CLASS

YES: ABILITY-GROUP IN THE FOLLOWING SUBJECTS FOR THIS CLASS: (CIRCLE AS MANY AS APPLY)

English..........	1
Reading..........	2
Mathematics......	3
Science..........	4
Social Studies...	5
Other:__________	6

---

6. Do you use a program of individualized instruction for any subject that you teach to the sampled class? (CIRCLE "YES" OR "NO".) If "yes," circle the subjects for which you use an individualized program of instruction for this class.

NO: DO NOT INDIVIDUALIZE INSTRUCTION FOR THIS CLASS

YES: USE INDIVIDUALIZED PROGRAM IN THESE SUBJECTS FOR THIS CLASS: (CIRCLE AS MANY AS APPLY)

English..........	1
Reading..........	2
Mathematics......	3
Science..........	4
Social Studies...	5
Other:__________	6

---

7. Do you meet with a team of other teachers who teach other academic subjects to the same group of students who are in the sampled class? (CIRCLE ONE NUMBER)

No, we do not have such a team (CIRCLE AND SKIP TO ...).... 1
Yes, we meet at regularly scheduled times.................. 2
Yes, we meet informally.................................... 3
There is a team, but we rarely meet........................ 4

8. Counting yourself, how many teachers are on your teaching team?

NUMBER OF TEACHERS ON TEACHING TEAM FOR THIS CLASS: _______

9. When you meet with your teaching team, how often do you do each of the following things? (CIRCLE ONE CHOICE FOR EACH ACTIVITY.)

a) spend the time grading papers from your own subject	OFTEN	SOMETIMES	SELDOM	NEVER
b) prepare your own lessons for your subject	OFTEN	SOMETIMES	SELDOM	NEVER
c) discuss the performance of individual students	OFTEN	SOMETIMES	SELDOM	NEVER
d) arrange to visit and observe another teacher with the same students	OFTEN	SOMETIMES	SELDOM	NEVER
e) plan curriculum so your subject-teaching is coordinated with the other teachers'	OFTEN	SOMETIMES	SELDOM	NEVER
f) arrange with another teacher to jointly teach the same class	OFTEN	SOMETIMES	SELDOM	NEVER

References

Alexander, W. M., & George, P. S. (1981). *The exemplary middle school.* New York: Holt, Rinehart, and Winston.

Ames, R., & Ames, C. (Eds.). (1984). *Student motivation.* New York: Academic Press.

Anderson, R.C., Hiebert, E.H., Scott, J.A., and Wilkinson, I.A.G. (1985). Becoming a Nation of Readers: The Report of the Commission on Reading. Washington, D.C.: National Institute of Education.

Baird, L. L. (1969). Big school, small school: A critical examination of the hypothesis. *Journal of Educational Psychology, 60,* 253-260.

Barker, R. G., & Gump, P. V. (1964). *Big school, small school: High school size and student behavior.* Stanford, CA: Stanford University Press.

Blyth, D. A., Hill, J. P., & Smyth, C. (1981). The influence of older adolescents: Do grade level arrangements make a difference in behaviors, attitudes and experiences? *Journal of Early Adolescence, 1,* 85-110.

Blyth, D. A., Simmons, R. G., & Bush, D. (1978). The transition into early adolence: A longitudinal comparison of youth in two educational contexts. *Sociology of Education, 51,* 149-162.

Bossert, S. T. (1979). *Task and social relationships in classrooms: A study of classroom organization and its consequences.* New York: Cambridge University Press.

Bossert, S. T., & Barnett, B. (1981). *Grouping for instruction.* San Francisco: Far West Laboratory for Educational Research and Development.

Bossert, S. T., Barnett, B. G., & Filby, N. N. (1984). Grouping and instructional organization. In P. L. Peterson, L. C. Wilkinson & M. Hallinan (Eds.), *The social context of instruction* (pp. 38-51). New York: Academic Press.

Boyer, E. L. (1983). *High school.* New York: Harper & Row.

Bridges, E. M., & Hallinan, M. T. (1978). Subunit size, work system independence and amployee absenteeism. Educational Administration Quarterly, 142, 24-42.

Brookover, W. B., Beamer, L., Efthim, H., Hathaway, D., Lezotte, L., Miller, S., Pasalacqua, J., & Tornatzky, L. (1982). Creating effective schools: An inservice program for enhancing school school learning climate and achievement. Holmes Beach, FL: Learning Publications.

Brophy, J. E. (1983). Classroom organization and management. Elementary School Journal, 83, 265-285.

Brophy, J. E. & Good, T. C. (in press). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), Handbook of research on teaching. (3rd ed.). New York: Macmillan.

Carroll, J. B. (1963). A model for school learning. Teachers College Record, 64, 723-733.

Coleman, J. S. (1980). How the young become adults. Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

Coleman. J. S., Campbell, E. Q., Hobson, C. J., McPartland, J. M., Mood, A., Weinfield, F. D., & York, R. L. (1966). Equality of educational opportunity. Washington, D. C.: U. S. Government Printing Office.

Coleman, J. S. et al. (1974). Youth: Transition to adulthood. Chicago: University of Chicago Press.

Doyle, W. (1984). Patterns of academic work in junior high school science, English, and mathematics classes. Paper presented at annual meeting of American Educational Research Association, New Orleans, April.

Doyle, W. (in press). Classroom organization and management. In M. C. Wittrock (Ed.), Handbook of research on teaching. (3rd ed.). New York: Macmillan.

Ebert, R. W., Kehoe, E., & Stone, J. A. (1984). The effect of school size on student outcomes. Eugene, OR: University of Oregon, Center ofr Educational Policy and Management (ED 245 382).

Epstein, J. L. (1983). Longitudinal effects of person-family-school interactions on student outcomes. In A. Kerckhoff (Ed.), *Research in sociology of education and socialization: Vol. 4.* Greenwich: JAI Press.

Epstein, J. L., & Karweit, N. (Eds.). (1983). *Friends in school.* New York: Academic Press.

Evertson, C. M. (1982). *Managing instruction for different class compositions.* Paper presented at the annual meeting of the American Educational Research Association, New York.

Farrar, E., Neufeld, B., & Miles, M. (1984). Effective schools programs in high schools: School promotion or movement by merit? *Phi Delta Kappan, 65,* 701-706.

Filby, N., Barnett, B. & Bossert, S. (1982). *Grouping practices and their consequences.* San Francisco: Far West Laboratory for Educational Research and Development.

Firestone, W., & Herriott, R. (1982). Prescriptions for effective elementary schools don't fit secondary schools. *Educational Leadership, 40,* 51-53.

Gage, N. L. (1984). What do we know about teacher effectiveness? *Phi Delta Kappan, 66,* 87-93.

Garbarino, J. (1978). The human ecology of school crime: A case for small schools. In E. Weeks & N. Harlow (Eds.), *School crime and disruption: Prevention models.* Washington, DC: National Institute of Education.

Gottfredson, D. C. (1985) School size and school disorder. Unpublished manuscript. Johns Hopkins University Center for Social Organization of Schools.

Gottfredson, G. D., & Gottfredson, D. C. (1985). *Victimization in six hundred schools: An analysis of the roots of disorder.* New York: Plenum.

Grabe, M. (1981). School size and the importance of school activities. *Adolescence, 16* 961), 21-31.

Karweit, N. (1981). Time in school. *Research in sociology of educati and socialization, 2,* 77-110.

Karweit, N. (1983). *Time-on-task: A research review* (Report No. 332). Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

Karweit, N. (1984). *Time-on-task reconsidered: Synthesis of research* on time and learning. *Educational Leadership, 41* (8), 32-35.

Kleinert, E. F. (1969). Effects of high school size on student activity participation. *NASSP Bulletin, 53,* 34-66.

Lipsitz, J. (1984). *Successful schools for young adolescents.* New Brunswick, NJ: Transaction Books.

McPartland, J. M. & McDill, E. L. (1977). Research on crime in schools. *Violence in schools: Perspectives, programs, and positions* (pp. 3-34). Lexington, MA: Lexington Books.

McPartland, J. M., & McDill, E. L. (1982). Control and differentiation in the structure of American education. *Sociology of Education, 55,* 77-88.

Messick, S., Beaton, A., & Lord, F. (1983). *National Assessment of Educational Progress reconsidered: A new design for a new era.* Princeton, NJ: National Assessment of Educational Progress.

National Center for Education Statistics (1984). *The Condition of Education: 1984 Edition.* Washington, D.C.: U.S. Government Printing Office.

Natriello, G., & Dornbusch, S. M. (1984). *Teacher evaluative standards and student effort.* New York: Longman.

Newmann, F. M. (1981). Reducing student alienation in high schools: Implications of theory. *Harvard Educational Review, 51,* 546-564.

Peterson, P. L., Wilkinson, L. C., & Hallinan, M. (1984). *The social context of instruction.* New York: Academic Press.

Purkey, S. C. & Smith, M. S. (1985). School reform: The district policy implications of the effective schools literature. *Elementary School Journal, 85,* 353-389.

Rosenshine, B. V., & Stevens, R. J. (1984). Classroom instruction in reading. Pp. 745-798 in P. D. Pearson (Ed.), *Handbook of reading research.* New York: Longman.

Slavin, R. E. (1983). *Cooperative learning.* New York: Longman.

Slavin, R. E. (1984). What we know about effective instruction. Baltimore: Johns Hopkins University, Center for Social Organization of Schools.

Slavin, R. E., & Karweit, N. L. (1982). *School organizational vs developmental effects on attendance among young adolescents.* Paper presented at the annual meeting of the American Pschological Association, Washington, D. C.

Slavin, R. E., & Karweit, N. L. (1984). Mastery learning and student teams: A factorial experiment in urban general mathematics classes. *American Educational Research Journal, 21,* 725-736.

Slavin, R. E., & Karweit, N. L. (1985). Effects of wholeclass, ability-grouped, and individualized instruction on mathematics achievement. *American Educational Research Journal.*

Wicker, A. W. (1969). Cognitive complexity, school size, and participation in school behavior settings: A test of the frequency of interaction hypothesis. *Journal of Educational Psychology, 60,* 200-203.

Winkler, D. R. (1980). Effects of sick leave policy and teacher absenteeism. *Industrial and Labor Relations Review, 2,* 232-240.

HARVARD UNIVERSITY
Graduate School of Education
Monroe C. Gutman Library, Appian Way
Cambridge, Massachusetts 02138

A HOUSE OF BRICKS

I told them a thousand times if I told them once:
Stop fooling around, I said, with straw and sticks;
They won't hold up; you're taking an awful chance.
Brick is the stuff to build with, solid bricks.
You want to be impractical, go ahead.
But just remember, I told them; wait and see.
You're making a big mistake. Awright, I said,
But when the wolf comes, don't come running to me.

The funny thing is, they didn't. There they sat,
One in his crummy yellow shack, and one
Under his roof of twigs, and the wolf ate
Them, hair and hide. Well, what is done is done.
But I'd been willing to help them, all along,
If only they'd once admitted they were wrong.
(Hay, The Builders, 1961)

Sixteen comprehensive reports on schooling in America have recently appeared, detailing the present woeful state of affairs, and offering a variety of solutions for falling test scores, increased functional illiteracy, poor teaching, undisciplined and disinterested children, inadequate leadership, and an incoherent smorgasbord of curricular offerings. Each researcher or team offers to build a new "house of bricks," offering more rigid, inflexible structures and strictures, "perks and punishments," most assuming as a given that the nineteenth century factory model of schooling will continue to survive unchanged.

Where is the library/media center in all of this paper blizzard of fresh ideas? When the indicators are that some 23 million adults are functionally illiterate, and that 13% of all 17 year olds in the United States are functionally illiterate (National Commission on Educational Excellence, A Nation at Risk, 1983) do any research organizations recognize the traditional role of libraries in stemming an avalanche of illiterati? Is there even a whisper hidden in the clamor of voices that says, "What about school libraries?"

Did a single researcher follow the students through their daily routines into the place that I normally regard as the information "soul" of every school, where learning activities and materials meet inquisitive minds in frenetic or leisurely embrace? The silence is almost total. Other than a single study that measured library space and amount spent on collections, relating these factors negatively to child development (Bloch, Effective Schools, 1983), not one analysis attended to libraries or media centers as part of educational assessment or as an instrument for educational improvement.

How did it happen? The star that rose so swiftly and shone so brightly, is wavering precipitously, in danger of obliteration. On every level, local, state, and particularly federal - at the same time as computers and computer literacy programs have captured the public in a panacea of purchase - library and media programs have lost the support of the power brokers and the general public. At the same moment that the possibilities of an information era are titillating imaginations, few have jiggled their minds sufficiently to realize that information storage and retrieval has always been intrinsic to libraries. Instead of selecting media specialists as superintendents and assistant superintendents for curriculum (they have both administrative and curriculum skills) their positions are being downgraded, while athletic directors or vocational education administrators continue a steady ascent to the top positions in educational administration. The enormity of changing a public image, so firmly ingrained even by librarians and media personnel themselves, in a world where access to information is truly the new power base, boggles the imagination.

That is not to say that NCES has neglected the gathering of statistics about the community of libraries - school, public, and academic. Numbers of volumes are reported annually by organizations and reported in the Digest of Education Statistics. This information is of limited value without knowledge of currency or research value, without any correspondence to the curriculum, without information about a range of materials - visual, oral, tactile, and their relatedness. A small independent school in New Hampshire proudly boasts of 10,000 volumes. Close examination revealed (Miller, unpublished report, 1984) that the volumes were primarily donations from well-wishers, and a spot-check of the shelves showed no use by students of the entire collection. In the decade that lies ahead access to information resources will be of far greater concern than quantities of bound volumes. When third grade children, particularly in middle-class communities, can do encyclopedia research via modem there must be a corollary use considered in statistics gathering. How many databases are being accessed by elementary/secondary/college/vocational students? How are the resources organized for maximum efficiency? What are the delivery systems? What are the programs in institutions to train teachers and students about the use and limits of the avalanche of information, soon to be available with a few strokes of the fingertips? How often can children access these information resources? Is it once a week for fifteen minutes in a weekly library trip. Recent information about computers (Becker, School Use of Microcomputers, 1984) indicates that the average student can use a microcomputer for 15 minutes per week. The use of the library in many schools, elementary and secondary, is also severely restricted, not for equipment or space but by structural limitations of school organization. Even counting numbers of things like books or filmstrips poses definitional problems (volumes or titles? purchased or acquired through grants and donations? useful or worthless?) The difficulties in keeping track of information about students is compounded when discussing the resources for learning. Definitional problems and lack of specificity plague ESEA Title III Offices in State Departments of Education, responsible for record gathering. Limited staffs, unable to undertake spot-checking, and schools with conflicting interests in determining future formula allocations have much to lose in making accurate reports. Secondary reporting sources compound the effects of inaccuracies and self-interest.

Methodology for both counting and reporting of information does not use current technologies. Sometime during the next decade it will become apparent that it is feasible to have the local organization report by modem into a clearly defined, standardized set of categories, updated regularly at the primary source. Unless such data become removed from identifiable funding decisions, reliability will continue to be suspect. Even the Harvard Annual Report by the University Librarian has much difficulty with the statistical aspects. Other than the accumulation of computerized data, a creative tour de force would best describe the individual faculty library reports. Each librarian is well aware of the effect of straying beyond prescribed limits. With regard to school instructional materials, their dating, relevancy, and use are more important data for researchers than present statistics.

NCES also publishes information about numbers of professional library/media personnel. Once again definitions and specificity are the issues. What is a library facility? Today, in the same way as grade and age groupings are difficult with disparate school organizational groupings, so too are libraries, media centers, instructional material facilities, teacher centers, television studios, computer operations or labs, equipment storage facilities, materials' warehouses all loosely joined together without any coherent or comprehensive definition. People who operate any of these facilities might be listed as media professionals for statistical purposes. When operations are expanded by schools to encompass the whole range of information access and retrieval, then a new nomenclature will replace the jumble that presently exists. If a distributed form of education begins to replace traditional class structure, and ancilliary organizations (i.e., museums, public libraries, homes) become primary learning sites, data gathering can be built into remote learning systems, with information available about student, teacher, and type of program.

For the past several years Gutman Library, Harvard Graduate School of Education, funded by a federal program, has been collecting information from 10,000 educational institutions about their use of microcomputers. The data are being mounted on Compuserve, a large national information utility. As schools begin to report numbers of microcomputers, conflicting information comes from individual schools and central administrative offices. Much of this equipment did not come from traditional budget sources. There are few accurate records. The lack of reliability is apparent in all data gathering around microcomputers by NCES, Market Data Retrieval, Gutman Library of Harvard Graduate School of Education, whether in the areas of hardware, software, or use of the computer. The limited knowledge acquired by central information sources leaves researchers wallowing in fuzzy figures and amorphous analytics. In a phenomenon, somewhat akin to the purchase of television receivers, that began as a bottom up movement, and was profoundly influenced by parents and the outside media, more substantive statistics would be available from equipment manufacturers and software producers. Flagrant violation of copyright by individual schools with regard to copying software will preclude any possible estimation of the numbers or types of instructional computer programs used in the nation's schools or homes. Even these constraints are negligible when compared to the lack of information by teachers about the kinds of computer use that engage their students. The submerged nature of the culture and the lack of instructional experience result in compilation of formalized programs by designated computer instructors. Research from the Harvard data base followed by visits to selected school systems reveal a standardized progression of computer use (Miller, A National Perspective-Microcomputers and Schools, unpublished speech, 1985).

Increasingly during the coming decade schools will begin to incorporate large-scale computerized curriculum systems on networked systems, supported by individual lap computers and enrichment or remedial materials to be used in the classroom or the home. These large systems will incorporate individualized information about each student as well as composite data. Now is the time for NCES to begin working cooperatively with schools in preparing meaningful data about student progress with new learning tools. Where are students using computers? Is there one computer in the classroom, a computer lab, take home computers, preparation of parents as well as teachers? Are teachers trained and involved in the planning process? Are the computers part of the regular curriculum or relegated to outside computer courses? How are they being used? What is the relationship to the reading or mathematics program? Are there supplemental materials? There is a need for contaxtual information. What applications programs will be part of the new curriculum? Data about the efficacy of technology has been of little apparent utility when the concentration is the equipment rather than the software or the usage.

As schools and universities concern themselves with access to computerization and wiring their buildings, much is taking place beyond the formal walls that will have a profound influence upon the future of education. The advent of telecommunications as part of a formal educational process is already in place in such institutions as Nova University, New York Institute of Technology, and New Jersey Institute of Technology. A kind of extended correspondence program or new form of bussing of information rather than students takes place nightly, with students involved in serious doctoral study. They sit in Arizona or Fort Lauderdale, studying and learning together. They are diverse ages, learning from home or place of business. They are part of a new world where jobs are not forever, and the need for new skills extends far beyond the traditional college cohort. Soon they will be joined by networks of gifted children in visionary projects out of Johns Hopkins or Pouch F, Alaska.

There are profound implications here for those who are planning the statistics gathering for the nation. These students involved in remote learning programs will extend the complexity of data gathering both generically and geographically. New schools will arise to be credentialled and incorporated into the network of independent or public institutions. Traditional organizations, with declining demographics, will seek students in corporations, among alumni, and in community groups. Home learning, already increasing nationally, will reach out to the preschool and the handicapped and the aged. New programs for delinquent youth and adolescent pregnant mothers are taking place outside the classroom or professional teaching faculty. Already the Reference staff at Gutman Library is reporting increased call, from doctoral students, for statistics on non-traditional learners and new technologies in the learning process. They want data on these populations, their ages, spending patterns, and educational backgrounds. They want to know more about changing careers and the need for increased schooling. They want better indexing and referencing of all information presently available. They want better use of current technologies in standard reporting.

As NCES begins its approach to the 21st century, the use of CD Rom disks for reporting of text information will play a part in the ability of research organizations and schools to more easily access and manipulate the mass of statistical data being collected. The organization and standardization of databases has begun to be of interest to multi-national corporations as they look to decision making based upon statistical analysis and examination of historical precedent. The same opportunities are available to organizations within the educational world if there is a real impetus for change, or if the whole search for educational excellence is not a charade or "a dance of legitimacy" (Deal, 1984).

The Public Library has a proud history in this country. It served to educate thousands of immigrants, and brought the culture of a new land and its language to peoples desperate for acculturation. Today the country still has great need for data, for information, and for knowledge. There are new sources. There are new possibilities. There are new methods for information transfer and retrieval. The library in school, university, corporation, and the community will continue to play an essential role. It will be the foundation for a new "house of bricks."

Inabeth Miller
Librarian to the Faculty of Education
Harvard Graduate School of Education
June 15, 1985

NOTES

1. Sara Kenderson Hay, "The Builders," in *Reflections on a Gift of Watermelon Pickle., and Other Modern Verse*. Stephen Dunning, et al, editors. (New York: Lothrop, Lee & Shepard Co., 1967).

2. National Commission on Excellence in Education, *A Nation at Risk: The Imperative for Educational Reform*. (Washington, D.C.: U.S. Dept. of Education, 1983): 8.

3. Alan W. Block, *Effective Schools: A Summary of Research*. (Arlington: Educational Research Service, Inc., 1983): 8-9.

4. W. Vance Grant and Thomas D. Snyder, *Digest of Education Statistics 1983-1984*. (Washington, D.C.: National Center for Education Statistics, 1985).

5. Inabeth Miller, report to Dublin School, Dublin, NH, February, 1985.

6. *School Uses of Microcomputers: Reports from a National Survey*, Issue No. 2. (Baltimore: Center for Social Organization of Schools, The Johns Hopkins University): 3.

7. Grant, *Digest of Education Statistics*: 49, 204.

8. Grant, *Digest of Education Statistics*: 200.

9. *Microcomputers in Schools, 1983-1984*. (Westport: Market Data Retrieval, 1985).

10. *National Database in Interactive Technologies*. (Cambridge: Gutman Library, Harvard Graduate School of Education, 1985). Online directory.

11. Inabeth Miller, *Microcomputers and Schools-a National Perspective*, speech to National Diffusion Network, March, 1985.

12. Terrence Deal, conversation with Inabeth Miller, May 1984.

PRIORITIES FOR FEDERAL EDUCATION STATISTICS

Richard J. Murnane

Graduate School of Education
Harvard University
Cambridge, MA 02138
Telephone: (617)495-3575

June 1985

350

INTRODUCTION

The purpose of this paper is to suggest priorities for the federal government's elementary and secondary education data collection efforts. My suggestions reflect not only my own ideas, but also things I have learned while participating in discussions of the National Research Council Committee on Indicators of Precollege Science and Mathematics Education. This committee has spent a great deal of time over the last six months discussing the quality of available data on U.S. elementary and secondary science and mathematics education. I have tried to indicate where my ideas differ from those voiced by other members of the Committee.

I have organized the paper in three sections: outputs, inputs, and private schools. In each case, I consider what we would like to know, what the available data are, and recommendations for the federal government's data collection efforts.

I. OUTPUTS

A. Test scores

1. National Assessment of Educational Progress

What is happening to the cognitive skill levels of children attending American schools? We know much more about this question than we did twenty years ago, primarily because of the National Assessment of Educational Progress (NAEP). The NAEP data have told us, for example, that:

--the reading skills of 9-year-old children improved over the 1970s, and the gap between the reading skills of black children and white children closed somewhat;

--the average science skills of students in all age groups fell between 1970 and 1973; the skill levels of 9- and 13-year-old students were stable over the period 1973-77, while the average science skills of 17-year-olds fell still further.

While the NAEP tests results have been informative, there are important questions concerning exactly what the tests measure. In particular, many analysts have argued that the NAEP tests do not measure higher order learning skills. Other analysts have argued that the tests do not even provide good measures of children's basic science literacy. The limitations of the NAEP tests and other tests of students' cognitive skills are worrisome for four related reasons.

First, we simply do not know whether the evidence on national trends in skill levels would be different if the tests provided better measures of cognitive skills, especially the critically important higher order skills.

Second, the lack of evidence on students' higher order skills makes it impossible to differentiate among alternative explanations for puzzles posed by the NAEP test results. For example, the NAEP results indicate that the reading skills of 9-year-olds improved over the 1970s, while the reading skills of 17-year-olds remained stagnant or fell. One possible explanation for this pattern is that the NAEP tests do not measure the true skill levels of older children as well as they measure those of younger children, and that our schools have in fact been as successful in educating older students as they have been in educating younger students. Another possibility is that the emphasis on basic skill acquisition in the early grades has had deleterious effects on students' acquisition of higher order reading skills. If true, this may have implications for how we teach children in the early grades. Yet another possibility is that many of the younger children benefited from participating in a well-developed Title I compensatory education program, while older children, if they participated in Title I at all, did so in Title I's uncertain, early years. We cannot differentiate among these possible explanations for the test score patterns until we have better measures of students' higher order reading skills.

Third, when test scores are used to assess the quality of educational programs, they tend to influence curriculum. The content of the NAEP tests may assume this role in the years to come as states contract with NAEP to provide detailed scores that can be used in statewide assessments of the quality of schooling. It would be extremely unfortunate if the lack of emphasis on higher order cognitive skills in assessment tests led to a reduction of emphasis on these skills in the curriculum.

A fourth reason for concern about test quality is that student test scores are the measure of teaching effectiveness used in almost all studies of the characteristics of effective teachers and the determinants of effective teaching. If the tests do not measure well the skills that chidren need to learn and that good teachers strive to teach, studies of the determinants of teaching effectiveness may give very misleading results.

Recommendations:

1. Continued funding for NAEP should have a very high priority. Current plans to increase the frequency with which math and science skills are tested should be retained.
2. The federal government should support Educational Testing Service's efforts to develop better NAEP tests. It is worthwhile not only to develop better multiple choice tests, but also to develop and utilize subtests that provide for open-ended responses to questions. This type of test item has greater potential for measuring students' higher order cognitive skills.
3. While it is critical to introduce better tests as soon as possible, it is important to retain enough of the old test items to permit comparison of new NAEP test results with the results of previous tests.

2. Scholastic Aptitude Tests (SAT) and American College Tests (ACT)

In recent years, comparisons among states of average SAT scores and average ACT scores have become increasingly popular. For example, they have a prominent place in the "Secretary's Wall Chart." It is well known that the average score in a state is sensitive to the average family income, the percentage of high school seniors in the state who take the test, and the percentage of students who attend private schools (Dynarsky, 1985; Howe, 1985). Moreover, in New Hampshire, the average SAT score is high, in part, because, included in the calculation are the high scores of students from other states who attend private high schools in New Hampshire. This is probably the case in other states as well. As a result of the influences on these nonschool variables on average SAT and ACT scores, these average scores are relatively poor indicators of the quality of public education provided to students in particular states.

Recommendation:

If the federal government must publish average SAT and ACT scores by state, publish alongside them an adjusted set of scores that takes into account the influences of participation rates, family income, and private school attendance. Such adjusted scores could be calculated relatively easily using multiple regression methods. My guess is that these adjusted scores would have a somewhat different pattern from the simple average scores. If this is the case, discussion of the reasons for the differences would be provocative, and maybe even informative.

3. International Association for Evaluation of Education Achievement (IEA) Cross-national Test Comparisons

Comparing, at one point in time, the average math and science scores of students in different countries poses a host of problems. In particular, differences in the quality of national school systems is only one of many reasons why average test scores differ among countries. Consequently, I am skeptical about the possibilities of drawing reliable inferences about U.S. education from international comparisons at a single point in time. Comparisons over time offer much better prospects, however. In particular, it is possible to examine how the achievement of U.S. students, as measured on the IEA tests, changes over time, and whether the position of U.S. students relative to students in other countries changes over time.

Recommendations:

1. Continue financial support for the IEA testing program, emphasizing the need to use test and sample designs that permit comparisons over time.
2. NCES should play a larger role in the implementation of the IEA tests in the United States.
3. Greater effort should be made to administer the IEA tests on a regular schedule so that comparisons over time can be made more reliably.
4. The Council of Chief State School Officers should be involved in administering the tests. This would improve local cooperation and reduce sampling bias due to nonresponse.

B. Dropout Rates

One important measure of the extent to which our schools accomplish the ambitious goal of educating all students is the percentage of students who graduate from high school. Most commonly, data are collected on dropout rates, which conceptually provide the same information as graduation rates. However, as Cook, Ginsberg, and Smith (1985) have documented, U.S. data on dropout rates (and graduation rates) are of very poor quality. The most common calculation method, comparing the number of students who graduate in year *n* with the number of students who entered high school in year n-3, is flawed for a number of reasons, the most important of which is the high mobility rate of American families. A critical effect of mobility is that it is not possible, using the standard calculation method, to distinguish a student who has left formal schooling entirely from a student who has transferred to a school in another jurisdiction. This problem is more severe the smaller the jurisdiction because mobility across jurisdictions is more prevalent. Consequently, dropout rates for individual schools, if calculated by the method described above, are probably less accurate than dropout rates for individual states--although even state dropout rates are influenced by family migration patterns.

It would be extremely valuable to have data series that provide comparable data on dropout rates for individual schools, school districts, and states. Such data are particularly important to have at this date as many states tighten requirements for high school graduation, requiring, for example, that students complete more math and science courses and pass a minimum competency exam. One of the adverse consequences of the new regulations may be that dropout rates increase. It is also likely that the effect of the new graduation requirements on dropout rates will be sensitive to the grade level at which minimum competency tests are administered and the extent to which systematic remedial help is available to students who fail. Reliable data permitting comparisons of dropout rates among states, and within states over time, would be valuable in determining how tightened requirements affect dropouts and whether the effects are sensitive to the details of the programs.

Recommendations:

1. NCES should work with the Council of Chief State School Officers (CCSSO) to develop and implement a uniform methodology for calculating dropout rates. Given the sensitivity of many state departments of education to federal pressure, achieving agreement will not be easy. I speculate that it is more than historical chance that explains the differences in methodologies used by individual states to calculate dropout rates. States probably have good reasons for choosing a particular methodology. Understanding why individual states calculate dropout rates as they do would be helpful in negotiating movement toward a uniform methodology. Consequently, a first step in improving data on dropouts is to systematically listen to the reasons dropout rates are calculated as they are in the individual states. The Council of Chief State School Officers may be an important vehicle for soliciting information on methodologies for calculating dropout rates, and for achieving agreement on a uniform methodology.
2. NCES should encourage, and if possible, fund studies that examine whether dropout rates as calculated by applying a new uniform methodology to school, school district, and state level data are close to dropout rates calculated from longitudinal data on individual students, such as that provided by HS&B. The reason is that following individual students over time is unquestionably the best way to learn about the dropout rates of students with particular characteristics who participate in particular kinds of

educational programs. In a world of no budgetary constraints, we would want all calculations of dropout rates to be done with individual, longitudinal data. If a new, common methodology for calculating dropout rates with aggregated data is indeed satisfactory, then the estimated rates should be similar to those calculated with data from HS&B.

C. Life Outcomes: Earned Income, Occupation, Probability of Employment

While American education has many goals, no one would deny that a central one is to prepare students with the skills and attitudes that will help them to earn a good living. How well do our schools accomplish this goal? This question has been hotly contested over the last 25 years, with advocates of the "human capital" approach documenting the accomplishments, while others, for example, Christopher Jencks (1972), documenting the failures. One point on which all analysts who have studied the "economic returns to education" issue agree is that learning more requires better data. The data sets used by Jencks and the human capital economists provide only minimal information about the kinds of education students received. We need datasets that provide detailed information on children's schooling as well as information on post-schooling careers and income paths.

Creation of the National Longitudinal Study of the High School Class of 1972 (NLS72) was a major step in creating a database that provided good information on students' school experiences and longitudinal information on their subsequent labor market experiences. An important limitation in the NLS72 database, however, is that students were interviewed for the first time during their senior year in high school, after most of the formal schooling was completed.

The High School and Beyond project is another important step in providing good longitudinal information on the schooling and subsequent labor market experiences of American students. From the HS&B data, we have already learned that the type of schooling students receive has a marked impact on their cognitive skills. I expect that subsequent research will tell us a great deal about the impact of particular types of schooling on subsecent labor market experiences. The longitudinal study slated to begin in 1988 (NELS) offers even more promise for increasing our understanding of the roles formal education plays in affecting life outcomes; NELS will conduct baseline interviews when children are still in elementary school (grade 8).

Recommendations:

1. Continue to fund additional follow-up surveys of both the NLS72 and HS&B cohorts. It is important to collect information on members of the NLS72 and HS&B samples as individuals age. Many important effects of different types of schooling may not become evident until individuals reach their mid-thirties.
2. Do not reduce the sample sizes in the NLS72 and HS&B follow-up surveys. This is critical because many important questions can be addressed with these data only if relatively large subsamples with particular characteristics are retained. For example, Manski (1985) has recently conducted an important

study of the characteristics of members of the NLS72 sample who became teachers. This could be done only because the overall sample size in the follow-up surveys was sufficiently large to include 510 individuals who became teachers. (See Section II for more on the Manski study.)

3. Provide sufficient funding for the NELS project to permit inclusion in the follow-up surveys of all members of the baseline sample, and to trace individuals who drop out of school or transfer from one school to another. Tracing students who transfer would permit an analysis of why students change schools, an important question that we know little about.

II. INPUTS

A. Teachers

1. Salaries

Common sense, as well as the results of research on the determinants of school effectiveness, point to the importance of teachers in the education process. The quality of the teachers in American schools depends on the career decisions millions of college graduates make about which occupation to enter, and how long to remain in that occupation. While many factors influence the attractiveness of alternative occupations, one critical factor is monetary compensation. For this reason it is important to collect annual data on the salaries of teachers relative to salaries in other occupations.

The evidence that relative salaries affect the career decisions of teachers is not overwhelming (to a large extent because the research is difficult to do), but evidence does exist. For example, studies have shown that teachers' mobility decisions depend on salaries (eg. Baugh and Stone, 1982). Comparison of student test score trends during the 1970s with teacher salary trends is also suggestive. Over the period 1970-1981, students' skill levels in the physical sciences fell much more dramatically than skill levels in biology did. During the years 1974-81 (the closest years for which I could find the relevant data), the pay premium for a college graduate trained in biology who took a job in business or industry instead of becoming a teacher grew from 12 percent of a beginning teacher's salary to 31 percent. The comparable pay premium for a graduate trained in physics grew from 39 percent to 86 percent (Bacharach et al., 1984, p. 66). These patterns suggest that one of the reasons for the test score decline in science was the increasing difficulty in attracting qualified science teachers, and that this problem was more acute for physical science teachers than for biology teachers.

There are numerous difficulties in compiling comparative salaries of teachers relative to salaries in other occupations. For example, should the 9 or 10 month salaries of teachers be inflated to make them comparable with the 11 or 12 month salaries in other professions? There is no consensus on the answer to this question, and the attractiveness of teaching salaries at any one point in time is sensitive to the decision. To my mind, however, the key value of relative salary data lies in comparisons over time. For example, given that the job of teaching is quite different from that of working in an industrial biology laboratory, we do not know whether college graduates trained in biology will find it more attractive to take a $15,000 job teaching for 10 months, or the 11 month biology job that pays 12 percent more. It is reasonable to assume, however, that schools are less able to attract talented biology teachers when the pay premium for working in industry grows to 31 percent.

Recommendation:

The federal government should publish on an annual basis comparisons of salaries in teaching with those in other occupations. The comparisons should be presented separately for each academic field. Useful comparisons would be starting salaries, and salaries for individuals with ten years of work experience. Data on starting salaries are collected currently by the Placement Center of Northwestern University, and are published by the National Educational Association. Consequently, it may not be necessary for NCES to do all of the data collection. In fact, it may be efficient to contract with Northwestern to collect comparable salary data for experienced workers. However accomplished, it is important that annual data be available to assess trends in the salaries of beginning teachers and experienced teachers relative to salaries in other occupations.

2. Quality of the Teaching Stock

Has the decline over the last 15 years in teaching salaries relative to salaries in other occupations led to a reduction in the quality of college graduates choosing to become teachers? The evidence on the average SAT scores of new teachers suggests that this is the case. As Vance and Schlechty (1982) have documented, these scores have fallen quite dramatically in recent years.

To my mind, it is important to collect data on the SAT scores of college graduates who enter teaching. The reason is that this gives an indication of how bright college graduates perceived the attractiveness of teaching relative to that of other occupations. I want to point out that several members of the Committee on Indicators of Precollege Science and Mathematics Education disagree with this suggestion. They argue that there is no evidence indicating that a teacher's SAT score is correlated with teaching effectiveness. Also, not all college graduates who enter teaching have taken the SAT. Consequently, it is not clear what the scores are telling us.

I believe that the criticisms of my fellow Committee members are important. On the other side, however, I think about the hard questions state legislators will ask Chief State School Officers in the coming years about whether the large increases in teacher salaries that many legislatures are passing have influenced the success of public schools in attracting talented college graduates to teaching. To my mind, SAT score information, if presented carefully, could help us in answering the legislators' question.

Recommendations:

1. NCES should attempt to collect information on the SAT scores of college graduates entering teaching. This effort will probably require significant collaboration with state teacher

certification agencies. It is important that the information pertain to college graduates who actually become teachers, rather than college freshmen who indicate that they plan to become teachers--these are quite different groups.

2. The statistics that should be reported are not the average SAT scores of graduates who become teachers, but rather the percentage of teachers who have SAT scores above a specific cutoff--for example, 450 on the verbal test. The reason is that the critical information the data conveys is the ability of the schools to attract literate college graduates. Percentage of teachers with scores above a cutoff point conveys this information more accurately than the average score of teachers does.
3. It is important to compare the percentage of teachers scoring above a cutoff point with the percentage of all individuals taking the test in that year who scored above the cutoff point. The reason is that the overall SAT score distribution changes from year to year, and it is necessary to compare teachers' scores with those of other students who took the test to judge how successful our schools have been in competing for talented college graduates in a particular year.
4. It is important to keep track of changes in the composition of the pool taking the SAT so that the effects of such changes on the test score distribution can be separated from changes in the ability of the public schools to attract talented college graduates.

3. Teacher mobility patterns and what they mean

The chart on page 27 of the NCES publication, Indicators of Education Status and Trends (1985), is provocative. It shows that the SAT scores of individuals who left teaching after a few years are higher on average than the scores of teachers who remained in the classroom. How should we interpret this information? Was this pattern different twenty years ago? Are there policy changes that might alter this pattern? Would higher salaries enable school districts to retain teachers with high academic ability? We simply do not know. It may be that teachers' decisions about how long to remain in the classroom are sensitive to salaries. It is also possible that many academically talented college graduates plan to teach for a few years, and then move on to a new challenge, such as law school, and that these plans would not be altered by moderate changes in teacher salaries.

To interpret descriptive patterns such as the one presented on page 27 of Indicators and to inform policy discussions about how to attract and retained talented college graduates into teaching, we need to learn more about the determinants of teachers' career decisions. We also need this information to improve the models that are used to predict teacher shortages and surpluses. To my knowledge, in no existing model of the teacher labor market (and I include the NCES model in this category) is the supply of new entrants or the turnover rate viewed as being sensitive to salaries. This is ironic in the sense that increases in teacher salaries are a common theme in the wave of current school reform legislation aimed at coping with the shortages predicted by the demand and supply models.

Why don't we know more about the determinants of teacher career patterns? One reason is that research on this set of issues requires data on the career decisions teachers make over time and on the attributes of their options. Little such data exist, although there are opportunities to create more at reasonable cost.

Recommendations:

1. Continue follow-up surveys of the NLS72 sample. Manski (1985) has identified 510 individuals in that sample who became teachers. One of the many benefits of following this sample as its members age is that we could study why some teachers remained in the classroom and others did not. It would be valuable in future follow-up surveys to include questions about the reasons for occupational changes, and about salaries before and after job changes.
2. NCES should ask Manski, Schlechty, and other researchers who have studied teachers' careers with the NLS72 data about how the HS&B and NELS follow-up surveys could be structured to overcome limitations of the NLS72 data for studying this set of issues.
3. The Current Population Survey (CPS) has a limited longitudinal component that can be used to examine the reasons teachers change jobs. Baugh and Stone (1982) have used CPS data to show that teachers' decisions about whether to change jobs are sensitive to salaries. My sense is that more could be learned from CPS data about the determinants of teachers' career decisions (as well as the career decisions of workers in other

occupations) if more attention to this set of questions was given in the design of questionnaire items. I encourage NCES to work with the Census Bureau in exploring the possibility of using the CPS surveys to learn more about teachers' career paths, and specifically about the reasons for job changes.

B. Capital Accounts

NCES collects a significant amount of information on the capital expenditures of local school districts. There is no question that the quality of the physical plant influences the quality of life for teachers and students in local school districts and that the cost of the physical plant is a significant burden on many school districts. There is a question in my mind, however, about the usefulness of the data that NCES collects on capital account expenditures. I have never seen a study that uses these data in a manner that improves our understanding of how U.S. education works.

Recommendation

NCES should explore whether the data it collects on capital account expenditures are used, and if so, whether the uses justify the cost of collecting and processing these data.

III. PRIVATE SCHOOLS

What roles do nonpublic schools play in educating American children? In recent years research by James Coleman (1982) and others using the HS&B data has increased our knowledge of the roles that certain types of nonpublic schools, especially Catholic schools, play. We still know very little about other nonpublic schools, however. I focus my attention on two types of nonpublic schools that may be playing an increasing role in American education: private schools as an after-hours complement to public schools, and for-profit private schools.

A. After-Hours Private Schools

A recent article in the New York Times reported that an increasing number of American children are attending private schools after regular school hours to supplement the instruction they receive in public schools. If this is indeed the case, then it introduces a new determinant of the skill distribution of American students, and another mechanism through which affluent families can provide for their children better education than that provided to children from poor families. (In Japan, private after-hours schools, called Juku, are an important mechanism through which middle class families prepare their children for the national exams that determine entrance to public universities; see Cummings, 1980 for more on Juku.) Little is known about after-hours private schools in the U.S., and it seems worthwhile to try to learn more about them.

Recommendations:

1. In the design of the new longitudinal study of American students (NELS), include a set of questions asking whether students do attend after-hours schools, and if so, what the schools do and what they cost.
2. Include a similar set of questions in the October CPS survey.

B. For-Profit Schools

The 1977 Census of Service Industries reported the existence of 2237 for-profit elementary and secondary schools in the U.S. These schools were very small, paid their teachers low salaries, and were disproportionately located in the South. A recent New York Times article reported that the number of for-profit schools is growing. The article described a number of relatively expensive for-profit schools serving students from upper middle class families. This description seems different from the very limited description of for-profit private schools that can be gleaned from the 1977 Census of Service Industries.

It would be worthwhile to learn more about the number of for-profit elementary and secondary schools in the U.S., where they are located, what tuitions they charge, and whom they serve. Once we know the answers to these questions, we could explore whether for-profit schools operate differently from not-for-profit schools. This would be valuable in thinking about the design of state regulations of private schools, and the design of voucher systems--topics of increasing interest in many states.

Recommendations:

1. Include questions in the next Census of Service Industries that will provide information about for-profit schools.
2. Explore whether it is possible to use IRS data to learn about trends in the number of for-profit schools, their locations, and their scale of operation.

IV. CONCLUDING COMMENT

I would like to express my support for the process NCES has initiated in attempting to improve the quality of federal statistics on elementary and secondary education. Asking a large number of individuals from different backgrounds for their ideas is essential to improving NCES's contribution to understanding U.S. education. It is also, however, an invitation to criticisms from many fronts. I admire the decision of the NCES leadership to solicit suggestions with the inevitable accompanying criticisms and the decision to make all of the suggestions public. I look forward to reading the suggestions of other commentators.

I conclude my comments with one final suggestion. As the NCES staff wades through the many sets of suggestions, it is inevitable that many suggestions cannot be implemented because of their cost, because of the politics of education in our federal system, and for a variety of other reasons. While most users of NCES statistics are acutely aware of their limitations, many are not aware of the reasons certain types of data are not collected. Producing a document that attemots to explain these reasons could be very informative. It is also possible that a statement of the reaons why potentially valuable data are not collected could lead to changes in the budgetary or political conditions that prevent the data collection.

REFERENCES

Bacharach, S.B., D.B. Lipsky, and J.B. Shedd. 1984. *Paying for Better Teaching: Merit Pay and Its Alternatives*. Ithica, N.Y.: Organizational Analysis and Practice.

Baugh, W.B. and J.S. Stone. 1982. "Mobility and Wage Equilibration in the Educator Labor Market." *Economics of Education Review* 2.

Coleman, J.S., T. Hoffer, and S. Kilgore. 1982. *High School Achievement*. New York: Basic Books.

Cook, C., A. Ginsburg, and M. Smith. 1985. "Researchers Find that Education Statistics Are in a Sorry State." *Basic Education* 29: 3-8.

Cummings, W.K. 1980. *Education and Equality in Japan*. Princeton, N.J.: Princeton University Press.

Dynarsky, M. 1985. "The Scholastic Aptitude Test: Participation and Performance." Working Paper No. 258. Department of Economics, University of California, Davis.

Howe, H. 1985. "Let's Have Another SAT Score Decline." *Phi Delta Kappan* 66: 599-602.

Indicators of Education Status and Trends. 1985. Washington, D.C.: U.S. Department of Education.

Jencks, C., M. Smith, H. Acland, M.J. Bane, D. Cohen, H. Gintis, B. Heyns. and S. Michelson. 1972. Inequality. New York: Basic Books.

Manski, C.F. 1985. "Academic Ability, Earnings, and The Decision to Become a Teacher: Evidence from the National Longitudinal Study of the High School Class of 1972." Working Paper No. 1539. Cambridge, MA: National Bureau of Economic Research.

Vance, V.S. and P.C. Schlechty. 1982. "The Distribution of Academic Ability in the Teaching Force: Policy Implications." Phi Delta Kappan 64: 22-27.

Products and Processes of the National Center for Education Statistics: An Agenda for the Next Decade

Gary Natriello
Teachers College, Columbia University

June 1985

Products and Processes of the National Center for Education Statistics: An Agenda for the Next Decade

Introduction

This review of the current data collection activities of the National Center for Education Statistics is divided into two major sections. The first section presents a set of eight major recommendations that apply to NCES plans in general. Although these general recommendations carry some implications for specific data collection activities, the emphasis is on the general needs for coverage in new content areas and the development of new processes to involve and serve users of NCES data. The second section of the paper reviews each NCES data collection activity and presents very specific suggestions. In some cases, the general arguments identified in the first section are illustrated by specific suggestions in the second section.

Section I: General Recommendations

The general recommendations for future NCES data collection activities are of two types. The first five recommendations deal with the content of current and new data collection activities. These recommendations typically concern educational phenomena that are likely to become increasingly important to educational policy makers in the years ahead. Other recommendations deal with the processes by which education statistics are collected. Such processes have an impact on both the quality of the data that is collected and on the utility of that data to potential users. These process recommendations imply a greater leadership role for NCES in organizing education data collection activities nationwide. They suggest not only a variety of ways in which NCES itself can collect data on educational processes, but also ways in which NCES can guide and coordinate the data gathering activities of others so that they can be employed in analyses of educational activities on a nationwide scale.

Recommendation 1: NCES Should Explicitly Consider Moving Beyond the Collection of Data on Schooling to the Collection of Data on Education

Most current NCES data collection activities focus on elementary and secondary schools. While schooling should remain at the core of NCES data collection plans, greater attention should be devoted to educational activities that extend beyond formal schools. Such greater attention is necessary for two reasons. First, non-school educational experiences are playing an important role in the lives of young Americans (Cremin, 1976; Leichter, 1975). Students are being exposed to a greater variety of educational phenomena than ever before from the mass media (Gans, 1967; Liebert, Neale, and Davidson, 1973; Lesser, 1974) to educational software and other new information technologies (Smith, 1982), to supplementary instruction in the growing number of propriety schools and tutoring services (Martin and McCartney, 1976). While we may not yet be in the age of Illich's (1971) learning networks, students are being exposed to a growing number of learning resources outside of the traditional schools. While the public schools are available to all and while we have spent considerable time and effort to attempt to secure equal educational opportunity, we know very little about the distribution of non-school educational resources.

A second reason for paying attention to the growing set of educational resources outside of traditional schools is to improve our understanding of the effects of schools themselves. If students and their parents are coming to rely increasingly upon non-school educational activities for the development of important skills and abilities, it will be impossible to understand the true effects of schools, public or private, without good information on the other educational resources available to students. Moreover, if more economically able parents secure these outside educational resources at higher rates there is a danger of attributing to ascriptive characteristics, effects that should be attributed to organized educational activities outside of public and private schools. Non-school educational activities may become important control variables, much like parent educational levels and economic resources, in understanding the effects of schooling.

Unlike public and private full-time day schools, non-school educational resources may be difficult to identify. Several strategies may be necessary to develop data on these educational activities. First, it would be useful to include items on regular NCES surveys of students and parents (e.g., High School and Beyond) which request respondents to report on the extent of their participation in non-school educational activities. A preliminary list of such activities would include: private lessons in music and/or art, private instruction in sports and other physical activities (e.g., tennis, horseback riding), participation in a computer users group, training related to a part-time job, attendance at an SAT or ACT test preparation course, remedial or supplementary instruction in one or more school courses, and training provided by a youth or community group such as the YMCA or the Boy Scouts or Girl Scouts. This strategy would permit analyses to determine the extent to which individual students participate in non-school educational activities and the impact of such participation on student performance in school classes and on standardized tests. Analyses could also be conducted to determine the access to such supplementary educational resources of students with different family and background characteristics.

A second strategy for collecting information on non-school educational activities would be to identify the population of educational service providers through state corporate records. Many supplementary educational activities are provided by private businesses, yet the education community and education policymakers have little current information on what may be the fastest growing part of the educational sector. A census of such organizations would begin to fill this gap in our knowledge.

A third strategy for collecting information on the extent of nonschool educational activities would be to identify a representative sample of communities and survey the available non-school educational programs available in the community. This strategy would permit analyses to determine the distribution of supplementary educational activities across communities with different demographic and economic characteristics.

The point of all of these data collection activities would be to begin to understand the extent to which non-school educational activities contribute to the development of elementary and secondary students in the United States. Educational policymakers will need to appreciate the potential and the pitfalls of such educational activities in developing a coherent approach to educational policy. The ready availability of such non-school educational options may suggest new policy initiatives. For example, recent plans to institute a tax credit or voucher system to enhance parental choice may be modified to permit parents who select a traditional public school for their children to use the credit to support supplementary educational activities. The ability to select individual activities as opposed to total school programs may give parents even greater choice and may further stimulate the growth of an educational services sector to meet rapidly changing educational needs.

Recommendation 2: NCES Should Conduct a Longitudinal Study of Students in the Elementary, Middle-School, and Junior High School Grades

The High School and Beyond data set has proven to be a valuable resource for investigators interested in understanding the progress of students from the high school to the work place. Longitudinal studies of the same sort are needed for younger children. Such studies are important for two reasons. First, since a number of problems associated with high school students (e.g., teenage pregnancy, dropping out, drug usuage) are now seen to have their roots in the years prior to high school (Furstenberg, 1976; Stroup and Robbins, 1972; Berg, 1980; Lipsitz, 1985), it is important to understand the processes operating earlier in the school career of students if we truly want to understand the problems of the high school years.

Second, there are a number of important issues that pertain directly to students in elementary, middle, and junior high schools. These include: the impact of the family on the early schooling experiences of children (Richardson, Galambos, Schulenberg, and Petersen, 1984), the impact of the various school-to-school transitions on young students (Simmons, Blyth, Van Cleave, and Bush, 1979), the effects of classroom organizational characteristics (Rosenholtz and Rosenholtz, 1981; Rosenholtz and Simpson, 1984), and the effects of the instructional and evaluative strategies adopted by teachers (Entwisle and Hayduk, 1981).

Mounting a longitudinal study of students from kindergarten through high school, while desirable, would be an expensive and difficult undertaking. If resource constraints make such a comprehensive effort impossible, a longitudinal study of junior high school students which follows them through the transition to the high school should receive top priority. Data from such a study could be put to good use in conjunction with the results of the High School and Beyond Study. A second study might be planned to follow students from the upper elementary or middle school grades into the junior high school. In both studies particular attention should be directed to the transition between educational organizations.

Recommendation 3: NCES Should Pay Greater Attention to the Collection of Data on School Processes

Notably absent from the High School and Beyond Study are data on school processes, the experiences of students and teachers inside of schools. While the study has good items on student background and experiences prior to high school, on student experiences outside of high school (e.g. TV viewing), and on student expectations for education and occupational attainment after high school (e.g., plans for post-secondary education), it contains virtually nothing about the processes that students currently experience in the school. The absence of these indicators severely limits the use to which educational researchers can put such a data set.

Social scientists (Sorensen and Hallinan, 1977; Barr and Dreeben, 1983) have recently directed attention to school processes, renewing an interest that characterized studies of schooling of the fifties and early sixties (Gordon, 1957; Coleman, 1961; McDill, 1969; McDill and Rigsby, 1973). Educators have placed additional emphasis on what schools can do to make a difference in the lives of students (Walberg, 1984). Yet, the High School and Beyond Study seems to treat the school itself like a black box into which student characteristics are poured and out of which student expectations and occupational experiences flow.

An effort should be made to identify key school processes and include indicators of such processes in future studies of schooling. Such processes might include: performance evaluation (Entwisle and Hayduk, 1981; Rosenholtz and Simpson, 1984; Natriello, 1982; Natriello, 1984; Natriello and Dornbusch, 1984), rules and procedures (McPartland and McDill, 1977; Gottfredson and Daiger, 1979), peer networks (Coleman, 1961; Hallinan and Tuma, 1978; Epstein and Karweit, 1983), group processes (Sharan, 1980; Sharan, Hare, Webb, and Hertz-Lazarowitz, 1980; Slavin, 1980), instructional strategy (Good, Ebmeier and Beckerman, 1978; Good, 1983), time on academic tasks (Karweit, 1983; Karweit, 1984), and standards for performance (Natriello and Dornbusch, 1984; Natriello and McDill, forthcoming). Recent surveys such as the National Institute of Education's Safe School Study (National Institute of Education, 1977) and Goodlad's A Study of Schooling (1983) might be used as initial sources for items relating to school processes.

Recommendation 4: NCES Should Organize the Systematic Collection of Data on Student Performance on Employer Sponsored Tests.

One ultimate outcome of schooling processes is student performance in the work place. From time to time there have been various reports of how recent graduates have fared on the performance tests administered by major employers. At times employers have pointed to the lack of student proficiency in basic communication and computation skills. Such charges are often said to be based on student performance on standardized tests used by major employers to screen potential employees. Closer analyses of the reactions of major employers to recent graduates (Levine, 1984) have suggested that the major deficiency of employees lies not in the area of basic skills but in the area of attitudes toward work and deportment.

In any case, employers are increasingly involved in collecting data on the performance of recent high school graduates (Committee on Ability Testing, 1982, chapter 4). NCES might capitalize on this trend by organizing and aggregating this data on the performance of recent graduates from a representative sample of major employers. Major sources of such data include the armed forces (The Armed Services Vocational Aptitude Battery - ASVAB - is the most used employment test.), the civil service, state and local governments, private sector employers, and various professional associations. Thus, performance data on recent graduates is being collected by a large number of employers across the nation. What remains to be done is the assembling of a carefully drawn sample of such information and systematic tracking and interpreting of the results of such tests over time. NCES might organize such an effort.

Recommendation 5: NCES Should Devote Greater Attention to those Educational Organizations that Serve Pre-school Youngsters.

While NCES currently collects data on the universe of public elementary and secondary schools with sample studies of private schools, it devotes relatively little attention to institutions serving the pre-school youngster. A notable exception is the survey of pre-primary enrollments using census data. Efforts should be made to work with appropriate agencies in the various states to collect data on the population of child care and early education institutions serving youngsters from birth to kindergarten. We have relatively little systematic national information on this rapidly growing sector of educational institutions. Such data would appear to be important for two reasons. First, in recent years there has been increased concern about the basic physical safety of children in such institutions. This has given rise to a more general concern for the overall quality of these programs. Second, we have a great educational opportunity that may be missed if such institutions are not identified and encouraged to promote the educational growth of young children. States should be encouraged to monitor such institutions and develop appropriate policies to insure that their educational potential is realized. As more and more women enter the work force full time, the role of such institutions in the U.S. is likely to become more important.

Recommendation 6: NCES Should Put in Place a Process for Ensuring that Its Large-Scale Data Collection Efforts Can be Linked with More Micro-level Studies

In the foreseeable future it is likely that only the federal government will be able to mount educational research projects involving the collection of large nationally representative data sets. Yet many new and interesting theoretical ideas and most richly textured studies of educational phenomena are developed by individual investigators or small teams of investigators working in a small number of schools with severely limited research budgets. If the large-scale data collection efforts of NCES are to profit from and enrich the work of most educational researchers, NCES will have to put in place a process to ensure that linkages can be made between its macro-level data collection program and micro-level investigations.

While a variety of strategies can be employed to establish and maintain this link, the following would seem to be the minimum required:

1) Develop clear and widely announced procedures for submitting items for NCES surveys.

2) In considering items for inclusion in NCES data collection efforts, those items which have been used successfully in small scale studies should be given priority over those

items which have not been proven in small scale studies.

3) NCES should convene a panel to develop a set of critical indicators of each level of schooling that might be used to describe school sites employed in small scale studies so that the educational research community can develop a clearer sense of the nature of the sites in which small scale studies are conducted.

4) NCES should work closely with the appropriate National Institute of Education Research and Development Centers and Regional Educational Laboratories to coordinate nationwide surveys with the on-going work of these major government sponsored, educational research institutions.

Recommendation 7: NCES Should Work with A Few States to Develop Pilot State Level Data Bases Related to Significant Educational Issues

It is becomming increasingly clear that the state is the level at which much educational policy-making takes place and will continue to take place in the future. With this in mind NCES should take a leadership role in developing model state level data bases in areas of significant educational concern. For example, the current interest in the calibre of the teaching corps is a crucial issue in education, and one that is not likely to go away. NCES has responded by developing a survey of teacher demand and shortage using a nationally representative sample of LEA's and private schools. While NCES should be commended for responding to increasing concern about the teaching force, the chief policy options to improve the teaching force are at the state level (United States Department of Education, 1984). Therefore, it would be far more appropriate for NCES to work with a few individual states to assist them in developing state level data bases relevant to this issue.

A two stage process should be initiated. In the first stage NCES would identify several states interested in developing a state level data base relevant to state policy making. NCES would then work with these SEA's to develop the data gathering procedures. In the second stage NCES might select the most successful data base design and use it as the model for a national data base assembled from data collected by individual states. Other SEA's would be asked to develop the designated state level data bases.

There are several advantages to such an approach. First, state level data bases would insure that the data gathered is at the appropriate level of aggregation for policy making. Second, NCES could provide a leadership role in helping state departments of education to develop the capacity for collecting and interpreting educational data. Third, NCES could develop national data sets, with some assurance that the actual data collection was useful to the data collectors (SEA's) and that the data collection was done properly.

Recommendation 8: NCES Should Maintain an Awareness of Emerging Technologies for the Transmission and Communication of Data to Relevant Groups of Users.

While the program of publications and computer data tapes now used by NCES has met the needs of many data users, NCES should maintain an awareness of new technologies for data transmission and communication with relevant users of its products. Although it is impossible to anticipate future developments in this area, several currently available technologies might offer increased access and ease of use for NCES data. These include:

1) Establishment of an on-line query-based information system to access NCES data. Such a system could be developed independently or in conjunction with a general information utility such as The Source or Compuserve. There is already evidence of the growing use of such information services by educational researchers and the general public (Pierce and Cooley, 1985).

2) The production of NCES data in new forms such as floppy disks for use on microcomputers. Microcomputers are moving quickly into universities, SEAs and LEAs, and educational policy agencies. NCES might increase the utility of much of its data to a wide variety of users by making it available in micro-readable formats.

3) NCES should develop a comprehensive program to make the research community more aware of its activities in the collection and processing of education data. The program of dissemination activities and curriculum materials developed for the 1980 Census provides some good examples of techniques that could be used by NCES.

Section II: Specific NCES Activities

The comments in this section regard specific NCES data collection activities. All of these comments, both those pertaining to existing items on NCES surveys and those recommended new items, are based on the package of NCES materials provided for this review. They do not reflect any other NCES activities.

I. Common Core of Data (CCD)

The Common Core of Data seems adequate to provide basic information on the universe of public schools and school districts in the United States. I have only three suggestions in this area.

First, the Local Education Agency Nonfiscal Report currently requests a report on only those programs that are part of the regular school system and that are financed by the local education agency or other public education unit. In order to understand the extent to which the public schools are involved (even without providing financial support) with a variety of new educational services such as day-care for young children or extended day programs for latch-key children, it would be important to request information on programs affiliated with each public school, even those which simply use the physical facilities of the public schools.

Second, respondents are allowed to designate the beginning of the range of the educational program as pre-kindergarten or kindergarten. It would provide more useful data on educational programs for very young children to have designations related to the age of the child.

Third, the current procedure for calculating student membership assumes that names of students on the rolls are automatically dropped after a specified number of days of consecutive absence. Yet there is no standard specification of the number of days after which students should be dropped. As a result, different district level practices may result in quite different reported enrollment levels. NCES should adopt a standard for retaining students on the rolls.

II. Sample Surveys

The comments on the NCES program of sample surveys pertain, for the most part, to specific survey items. These comments are organized in terms of the six current surveys conducted by NCES.

Private School Survey

The following items on the private school survey seem worthy of additional attention:

a) On page 6, question 14, respondents might be asked to note the amount of funds available for financial aid during the current academic year.

b) On page 7, question 15, respondents might be asked to report on programs offered primarily to extend the school day and used by working parents for child care purposes.

c) On page 10, question 19, respondents might be asked whether this school restricts admission on the basis of membership in a religious, cultural, or language group. If the answer is yes, they might be asked to note the nature of the restriction.

d) On page 10, respondents might be asked to report the percentage of initial applicants to this school that are finally admitted.

e) Respondents should be asked to report the name of the public school or schools to which students in this school would normally be assigned. This information can then be further identified with the NCES school code. This item might be included in a future survey of parents.

Public School Survey

Administrator Survey

a) It would be useful to include a set of items on this survey that parallel those on the private school survey. For example, the questions on incentive systems could be included in both surveys.

b) Question 23 on page 4 should be deleted. It asks for the administrator's opinion regarding curriculum requirements when no information on current requirements is requested.

Teacher Survey

a) In general this survey solicits information about school-related activities (e.g., question 17, page 5) but never inquires about instructional activities. We have no idea what kind of instruction is going on in classrooms. Unfortunately, this continues an NCES pattern of ignoring instructional processes in classrooms.

b) Questions 36 solicits the teacher's opinion on the number of years of study that should be required without determining current school requirements.

Recent Survey of College Graduates

a) If not already done, information on the college attended (e.g., tuition rate, etc.) should be collected and matched with student responses.

b) Information on the teacher certification requirements of the state in which these students are certified should be collected and matched with student reports. This would make it possible to examine the impact of new state certification policies.

Survey of Teacher Demand and Shortage

a) As noted earlier, this information would be more appropriately collected on the state level. Any national representative sample should also be representative for each state so that SEAs can use the results for planning purposes.

High School and Beyond

Sophomore First Follow-Up

a) More detailed questions regarding curriculum choice might be added as a follow-up to question 4 on page 3. Students should be asked to report why they decided to take a course or not to take a course. See Garet and DeLany (1985).

b) Questions should be asked about experiences in particular classrooms. See Goodlad (1983) and Natriello and Dornbusch (1984) for examples.

c) Questions 90, 91, and 92 should be deleted. The information on anticipated expenses will be of doubtful value.

d) Question 128might contain a response category for special college funds established by parents.

Senior Year First Follow-Up

a) On question 11 on page 10, response categories should permit a distinction between getting fired and other reasons for leaving a job.

b) On question 37 on page 22, the distinction between course credits and semester or quarter credits is unclear.

In General

a) An attempt should be made to develop a broader range of outcome tests. The currents tests represent a narrow range of academic skills.

b) Questions should be added pertaining to outside supplementary instruction while in high school (e.g. SAT tutoring, music lessons, use of SAT prep software, vocational training etc.).

Library and Media Center Survey

a) A question should be added to request the costs for the maintenance of computers on page 4.

b) Questions should be added to determine the patterns of current use and future needs for various types of data bases.

III. Other Agency Data

Preprimary Enrollments of Children 3-5 years old

a) This is a good effort to develop data on this important group. Additional efforts would be useful, perhaps dealing with educational options for even younger children.

References

Barr, R. and R. Dreeben. How Schools Work. Chicago: University of Chicago Press, 1983.

Berg, I. School Refusal in Early Adolescence. Pp. 231-249 in L. Hersov and I. Berg (eds.) Out of School: Modern Perspectives in Truancy and School Refusal. New York: Wiley, 1980.

Coleman, J. The Adolescent Society. New York: Free Press, 1961.

Committee on Ability Testing, Assembly of Behavioral and Social Sciences, National Research Council. Ability Testing: Uses, Consequences, and Controversies. Part 1: Report of the Committee. Washington, D.C.: National Academy Press, 1982.

Cremin, Lawrence A. Public Education. New York: Basic, 1976.

Entwisle, Doris R. and Leslie Alec Hayduk. Academic Expectations and the School Attainment of Young Children. Sociology of Education, 54, 1981:34-50.

Epstein, J.L. and N. Karweit (eds.). Friends in School: Patterns of Selection and Influence in Secondary Schools. New York: Academic Press, 1983.

Furstenberg, F. Jr. Unplanned Parenthood: The Social Consequences of Teenage Childbearing. New York: Free Press, 1976.

Gans, Herbert J. "The Mass Media as an Educational Institution," Television Quarterly, 6, Spring 1967:2037.

Garet, Michael S. and Brian DeLany. Decisions and Random Draws: Course Choice in Math and Science. Stanford, California: School of Education, Stanford University, 1985.

Good, T. Research on Classroom Teaching. In L.S. Shulman and G. Sykes (eds.), Handbook of Teaching and Policy. New York: Longman, 1983: 42-82.

Good, T., H. Edmeier, and T. Beckerman. Teaching Mathematics in High and Low SES Classrooms: An Empirical Comparison. Journal of Teacher Education, 29, 1978:58-90

Goodlad, J. A Place Called School: Prospects for the Future. New York: McGraw-Hill, 1983.

Gordon, C. Wayne. The Social System of the High School. Glencoe, Illinois: The Free Press, 1957.

Gottfredson, G. and D. Daiger. Disruption in Six Hundred Schools. Report No. 289. Baltimore, Maryland: Center for the Social Organization of Schools, The Johns Hopkins University, 1979.

Hallinan, M.T. and N.B. Tuma. Classroom Effects on Change in Children's Friendships. Sociology of Education, 51, 1978:170-282.

Illich, Ivan. Deschooling Society. New York: Harper and Row, 1971.

Karweit, N. Time-on-Task: A Research Review. Report No. 332. Baltimore, MD: Center for the Social Organization of Schools, Johns Hopkins University, 1983.

Karweit, N. Time-on-task reconsidered: a synthesis of research on time and learning. Educational Leadership, 41, 1984:33-35.

Leichter, Hope Jensen, (ed.). The Family as Educator. New York: Teachers College Press, 1975.

Lesser, Gerald S. Children and Television: Lessons from Sesame Street. New York: Random House, 1974.

Levine, Marsha. Summary of Report: Survey of Employer Needs. Washington, D.C.: Committee for Economic Development, 1984.

Liebert, Robert M., John M. Neale, and Emily S. Davidson. The Early Window: Effects of Television on Children and Youth. New York: Pergamon Press, 1973.

Lipsitz, Joan. A Statistical Portrait of American Adolescents. Common Focus: An Exchange of Information About Early Adolescence, 6, 1985:1-5.

Martin, R. and R.J. McCartney. Education's Big Boom is Ending, but Studies to Get More Diverse. Wall Street Journal, April 8, 1976.

McDill, E.L. and L.C. Rigsby. Structure and Process in Secondary Schools: The Impact of Educational Climates. Baltimore, The Johns Hopkins University Press, 1973.

McDill, E.L., L.C. Rigsby, and E.D. Meyers, Jr. Educational climates of high schools: their effects and sources. American Journal of Sociology, 74, 1969:567-586.

McPartland, J. and E.L. McDill, (eds.), Violence in Schools. Lexington, Mass.: D.C. Heath and Company, 1977.

National Institute of Education. Violent Schools-Safe Schools: The Safe School Study Report to the Congress. Washington, D.C.: U.S. Government Printing Office, 1978.

Natriello, G. Organizational Evaluation Processes and Student Disengagement in Secondary Schools. St. Louis: Final Report to the National Institute of Education, 1982.

Natriello, G. Problems in the Evaluation of Students and Student Disengagement from Secondary Schools. Journal of Research and Development in Education, 17, 1984:14-24.

Natriello, G. and S.M. Dornbusch. Teacher Evaluative Standards and Student Effort. New York: Longman, 1984.

Natriello, G. and E.L. McDill. Performance Standards, Student Effort on Homework and Academic Achievement. Sociology of Education, forthcoming.

Pierce, Jean and William W. Cooley. A Communications Network for Educational Researchers. Educational Researcher, 14, 1985:14-17.

Richardson, Rhonda A., Nancy L. Galambos, John E. Schulenberg, and Anne C. Petersen. Young Adolescents' Perceptions of the Family Environment. Journal of Early Adolescence, 4, 1984:131-153.

Rosenholtz, S.J. and S.H. Rosenholtz. Classroom Organization and the Perception of Ability. Sociology of Education, 54, 1981:132-140.

Rosenholtz, S.J. and C. Simpson. The Formation of Ability: Developmental Trend or Social Construction? Review of Educational Research, 54, 1984:31-64.

Simmons, Roberta G., Dale A. Blyth, Edward F. Van Cleave and Diane Mitsch Bush. Entry into Early Adolescence: The Impact of School Structure, Puberty, and Early Dating on Self-Esteem. American Sociological Review, 44, 1979:948-967.

Sharon, S. Cooperative Learning in Small Groups: Recent Methods and Effects on Achievement, Attitudes, and Ethnic Relations. Review of Educational Research, 50, 1980:241-271.

Sharon, S., P. Hare, C.D. Webb, and R. Hertz-Lazarowitz, (eds.), Cooperation in Education. Provo, Utah: Brigham Young University Press, 1980.

Slavin, R. E. Cooperative Learning. Review of Educational Research, 50, 1980:315-342.

Smith, Anthony. Information Technology and the Myth of Abundance. Daedalus, 3, 1982:1-16.

Sorensen, A.B. and M.T. Hallinan. A reconceptualization of school effects. Sociology of Education, 50, 1977:273-289.

Stroup, A.L. and L.N. Robins. Research Notes: Elementary School Predictors of High School Dropout Among Black Males. Sociology of Education, 45, 1972:212-222.

United States Department of Education. The Nation Responds: Recent Efforts to Improve Education. Washington, D.C.: U.S. Government Printing Office, 1984.

Walberg, Herbert J. Improving the Productivity of America's Schools. Educational Leadership, May 1984:19-27.

FEDERAL COLLECTION OF SCHOOL FINANCE DATA: NEW NEEDS FOR AN ERA OF EDUCATION REFORM

By

Allan Odden

University of Southern California

In this paper I will address my comments primarily to data needs for analyzing and assessing - for policy purposes - the level, allocation and use of resources for elementary and secondary schools. I will suggest changes for data that have (or have not been) collected in the past to analyze the traditional fiscal issues in education and school finance, and will suggest a number of new types of data that are needed for analyzing issues related to education excellence concerns, which are unlikely to whither in the future.

As an overview, I will make the following general suggestions:

1. There is no question that the state is the primary actor in education policy, finance or otherwise; federal data collection should reflect this fact. Thus, data should be collected on a district and state basis; if a sample of district data are collected, which is appropriate for many data sets, the sample should be REPRESENTATIVE FOR EACH OF THE FIFTY STATES, as well as for the nation as a whole. Many fiscal data in the past have been used for state-by-state analyses even though sampling procedures technically did not produce a valid sample for each state.

2. School finance issues now should be framed as resource not just fiscal allocation issues, since the relationship between cost and quality has been linked finally by policy pressure. The policy interest is in improving schools; resources, which are broader than money, become means to those ends. The education excellence or reform movement has forced the merger of finance analysis with education effectiveness analysis, if not totally at least to a considerable degree. Thus revenue and expenditure data should be augemented by data on what they buy such as human resources (teachers and administrators), curriculum and time.

3. The policy interest in improving schools and the link between resource analysis and education effectiveness also raises the need for data on the allocation and use of resources: money, people, time and curriculum. To provide the analyses needed for policy questions, knowing how resources are used becomes as critical information as knowing the level of resources and how they are distributed across districts. Thus, more detailed information is needed on how revenues are used, including at least expenditures by program.

In short, federal data collection concerned with resources should identify the state (1) as the unit of analysis, should broaden

---

1. Analysis for federal policy targets the state as the unit but uses district data too; analysis for state policy requires district, school, classroom and student level data.

the definition of resources beyond revenues and expenditures to include people, time and curriculum, and should provide data on resource allocation and use within districts as well as resource distribution across districts.

The remainder of this paper has three sections. Section one identifies revenue and expenditure data that should be collected to conduct traditional school finance analyses which are still important, as well as some new fiscal analyses. Section two outlines data needs for human resources - teachers, administrators, other education personnell and students. Section three identifies data needs for time and curriculum.

1. REVENUE AND EXPENDITURE DATA

NCES should continue to collect state aggregate revenue and expenditure data, and revenue and expenditure data for a representative sample of districts, where the sample is representative of each of the fifty states as well as the nation as a whole. Even federal policy is now increasingly being conceived as related to and sometimes even an adjunct to state education policy; this means that federal data efforts need to have accurate data for each state, not just the nation as a whole. This probably would entail a shift in the sampling frame now used to obtain fiscal data.

For EXPENDITURES, there should be a strong push for states to collect data by program, so that NCES data can provide as much data as possible on how revenues are used. Program expenditure data become even more important in the context of state education policy interest and the structuring of new state categorical aid programs. Many states have created new education reform categorical programs - such as Florida's writing program, California's tenth grade counseling program, and South Carolina's and Texas' preschool programs for at-risk perschoolers. Other states, like Arkansas, have established new educational standards or raised high school graduation requirements and expect increased general revenue dollars to be spent to implement those objectives. Thus there is strong policy interest in how dollars are spent locally, i.e., on what programs dollars are spent.

Thus, NCES expenditure data should include:

o EXPENDITURES BY FUNCTION such as administration (board of education, central office with perhaps a breakout for curriculum units, categorical program administration, and building administration), instruction, transportation, operation and maintenance of plant, etc. Within the instructional category, expenditures on items such as classroom teachers, special program teachers, pupil support services, curriculum development and implementation, textbooks, materials and computer hardware and software (new categories) are needed. There still is a need to know expenditures by the traditional functions; indeed, a key interest overtime is in increasing instruction expenditures.

o EXPENDITURES BY PROGRAM such as regular education; programs for special need students such as the handicapped, low achieving (or economically disadvantaged), limited-English proficient and gifted and talented; by curriculum content area or at least broad categories of curriculum such as mathematics, science, social studies, language arts, etc.; and by elementary, middle and secondary level. Given current strong, nationwide interest in strengthening the high school curriculum, expenditures by broad categories of curriculum at the secondary level have a priority. While these data may need to be developed over a long time period, the fact is that policy interest at both the state and federal level need these data for policy analysis. My hunch is that state policy needs will provide the pressure to produce these data; federal interest in data collection of this sort, then, could give additional impetus to state efforts.

o at the least, expenditures for the general fund, restricted fund and capital fund should be gathered.

For REVENUES, NCES needs to expand the detail of data collected for both state and local revenues. Detailed data on federal education revenues have dominated the revenue side of federal data collection in the past; that detail may need to be retained for federal policy purposes. But many more subcategories of state, as well as local revenue need to be collected. Policy needs at both the state and federal level require data on state support for the general program, programs for special need student, as well as the host of rapidly increasing state education reform categoricals. The fact is that states are expanding the number of categorical programs through which funds are distributed, further restricting local discretion in resource use, but also stimulating districts to expend funds in support of education reform objectives, for which there is generally a healthy overlap between most states and the federal government at least as suggested in The Nation at Risk report. At the local level, the major new phenomenon is the expansion of non-property tax revenues which range from local education foundations to school-business partnerships and local option sales and income taxes. These revenues sources are of interest at the state and federal level.

Thus, NCES REVENUE data should include:

o some smaller sample of federal revenues

o state revenues by such categories as general aid, transportation, and facilities; programs for special need students such as special education, compensatory education, bilingual education, vocational education and gifted and talented; and education reform categoricals such as for high school graduation requirements, content areas such as science or writing, merit schools, tenth grade counseling, longer days and years, prekindergarten or day care at schools, full day kindergarten and mentor teacher, master teacher or career ladder programs, etc

o local revenues by such categories as local property taxes, local sales taxes, local income taxes; the traditional fees and community services; and new sources such as local education foundations, fee-for-service activities of the business sort such as new day care or preschool programs, private (either parent or corporate) contributions, and a miscellaneous category which in California would include urban development grants and developers fees.

In addition, local property wealth data (adjusted by differing local assessment practices), average district household income and local property, income or sales tax rates for schools are needed. I know such data are difficult to collect in many places. But gathering them for states for which they are available would be a good first step.

By using data from other governmental agencies, data on total state and local taxes, state and local taxes per capita and per $1000 of personal income and any indicators of the character of the incidence of state local taxes would be useful in NCES compilations. NCES compiles, at least to some degree, some of these data from the Census Bureau's periodically collected data sets..

Finally, everyone uses the annual and TIMELY NEA ESTIMATES OF SCHOOL STATISTICS, although this is an expensive activity to undertake (for the NEA). The organization or at least some people in it are not convinced, and rightly so in my opinion, that data collection of this sort is worth the costs. Yet, these data are used in all timely - at least for policy purposes - analyses of resource issues across the fifty states, and are a critical data set. I would suggest that NCES collect these data or give a subcontract to NEA to continue to collect these data. The data need to be produced quickly - estimates for the current academic year need to be available mid-way through the academic year. Thus, if NCES were to collect the data, the understanding would be that publication by March was mandatory; a subcontract to NEA might be an easier route to succeed on the timely criterion. And for these data, I do not believe the NEA label makes them suspect; the revised NEA estimates and the much later NCES figures usually differ for definitional and not accuracy reasons.

2. HUMAN RESOURCES DATA

More detailed data on human resources, i.e., the people in the education system - students, teachers and administrators - are needed not only because policy interest is forging a tighter connection between dollars and what they purchase, but also because interest in the human resource component of education is high. Indeed, many people feel that the key to success of current education reform is the ability of the country and each state to increase the number of high quality people entering and remaining in the teaching profession. A number of states, groups of deans from the top colleges of education and national groups, in fact, have decided that strengthening the teaching profession is the key to bringing about a restoration of the deteriorated excellence of the nation's schools. Finally, many states are restructuring the teaching profession by enacting career ladder, mentor teacher and master teacher programs; some states may create a new category of teacher, called the adjunct teacher - a professional in some field who teaches part-time but does not hold a teaching credential. Such a category would open the education system to high quality people without reducing standards for full entry into the teaching profession for people who want to teach full-time. In short, as the structre of the teaching profession changes, information on teachers, including types of teachers, becomes important. Finally, since teacher compensation is being changed to conform to the elements of the new structures, more information on teacher compensation also is needed.

Thus, NCES teacher and administrator data should include the following:

o numbers (FTE) of classroom teachers by teaching area (elementary, mathematics, science, language arts, social studies, art, music, etc.) and by special program (compensatory education, bilingual, special education, gifted and talented)

o numbers (FTE) of nonclassroom professionl personnell such as guidance counselors, psychologists, etc.

o numbers of teachers entering and leaving the above categories each year

o numbers (FTE) of teachers at different levels in a career ladder, numbers of master teachers, numbers of mentor teachers, or whatever categories states use for new structures. At some time, the data could be organized into a limited number of discrete, generic categories such as career ladders, master or mentor teachers, etc.

o numbers (FTE) of teachers working 10, 11 and 12 month contracts

o any measure of quality of those entering, remaining or leaving - score on a national teacher exam if one is given, perhaps the NTE until that time period; area in which degree earned (education, a discipline such as mathematics, science, English, history, etc); quality of college attended (using some list that ranks colleges and universities into a small number of quality levels); SAT or ACT score; or whatever.

o numbers of adjunct teachers, teachers with "emergency" or "waiver" credentials, paraprofessionals working in the classroom, etc.

o numbers (FTE) of administrators - central office line staff, curriculum coordinators and central office curriculum staff, staff development, building principals, assistant principals, assistant principals for instruction.

The idea would be to begin identifying new categories within which to group teachers or in which states are now grouping teachers, show where those categories are being used and how many new teachers are in each one. The data might be fragmentary for awhile, but would show the evolving nature and structure of the teaching profession. Fragmentary data are acceptable when reported on a state-by-state basis. The need for data on the evolving structure of the teaching profession highlights the need to focus on the state level.

NCES data for teacher compensation expenditures would include the following:

o average, median, minimum and maximum beginning salary

o average and median salary

o average, median, minimum and maximum top salary

o numbers of teachers by years of experience as well as by educational attainment - M.A., Ed.D., Ph.D.

o total spent for base salaries, career ladders (total for each ladder), mentor teacher, master teacher or merit teacher programs o total for benefits, and amount for for each benefit - pension, health, dental, life insurance, etc.

o similar data for administrators and noncertified personnell.

Student data suggestions are given here in the human resource section of the paper. Many of these student data are now collected; I list them only to emphasize that their collection should continue. NCES STUDENT data should include:

o Total FTE, ADA and ADM (although an FTE figure would suffice) and by elementary, middle and secondary level

o Total FTE by ethnic background (Anglo, Black, the categories of Hispanic, Asian American, Native American, etc.) and in special programs - compensatory education, bilingual education, special education, gifted and talented.

Enrollment projections for the nation and if possible for each state should be given every two to three years; most people were surprised by declining enrollments in the 1970s. If the lesson then was to watch trends, the lesson soon was forgotten; most have been surprised by increasing enrollments in the 1980s. Some states produce enrollment projections; others do not. Highlighting enrollment trends nationally at least would keep trends in the numbers and types of students more in the public eye.

For all these human resource data, I am probably suggesting an expansion of the common core data now collected each fall; the new data I have suggested for teachers is critical since significant changes in the structure of the profession are evolving, turnover of those now in teaching will be substantial in the next ten years due to retirements, and newly hired teachers will be numerous due to teacher retirements and rising enrollments in the next ten years. Collecting ad analyzing these data will be crucial for education policy interests at least over the next decade and a half.

3. TIME AND CURRICULUM RESOURCES DATA

Time and curriculum are the other two categories of resources into which revenues are converted, and both time and curriculum content are areas of high interest for most education excellence reforms. Thus, some aggregate data on these two variables are needed, even though variations in time and curriculum for individual students over time are the key data needed for analyzing the true impacts of these variables. Nevertheless, the macro-context of time and curriculum issues also is important. In addition, no organization systematically collects the type of data mentioned below and reports it periodically.

Thus, NCES TIME data should include:

o length of school year and day for teachers and for students by elementary and secondary level, preferrably in days and minutes.

o allocated times for instruction, by some categories of content areas (mathematics, language arts and science at the elementary level, and academic areas such as mathematics, science, language arts and social studies at the secondary level)

NCES CURRICULUM data should include:

o high school graduation requirements adjusted to some common norm across states

o numbers of secondary student course sections (adjusted to some standard such as number of sections per 1000 students and adjusted for enrollment fluctuations) by academic areas (mathematics, science, language arts, social studies, fine arts, performing arts, etc.) and numbers of specific courses within academic areas such as algebra I, algebra II, geometry, calculus, general mathematics, etc. for mathematics.

SUMMARY

In sum, future NCES collection of school finance data should have the following characteristics:

o focused on the state as the unit of analysis - sample district data should produce a valid sample for each state

o broadened to include resource and not just fiscal data now that policy issues have linked school finance with school effectiveness

o additional detail on state and local revenue

o expanded information on expenditures such as expenditures by program

o expanded detail on teachers - numbers by type (career, mentor, master, etc.), quality entering and remaining in the profession, and expenditures for base salaries as well as increments for career ladders, mentor or master teacher programs

o additional information on time and curriculum.

The Elementary/Secondary Redesign Project: Assessing the Condition of Education in the Next Decade

Penelope L. Peterson
University of Wisconsin-Madison

The stated purpose of the National Center for Education Statistics (NCES) is to collect data on the condition of education in the United States and to publish reports analyzing and interpreting these data (National Center for Education Statistics, 1984). One way to conceive of the role of the NCES is that the Center should collect data and provide interpretations of the data that are sufficient to give a "reading" of the general health of the nation's educational system. Just as a physician uses a few vital signs to assess the general health of the human organism, NCES should focus their efforts on a few selected areas of education rather than attempt to collect extensive data on a large number of variables. Thus, rather than collecting additional data on many new variables, I propose that NCES collect new and additional information in three major areas related to elementary and secondary education. Assessing these three major areas might be compared to taking the pulse, measuring the blood pressure, and examining the reflexes of the American educational system. In the sections that follow, I provide a brief discussion and rationale for the data collection in each of these new areas of education.

Three Needed Areas of New Data Collection

To assess the condition of American education in the next decade, we need information that addresses three major questions: (1) What are students doing and learning in the nation's schools? (2) What are the concerns and stresses facing teachers in the nation's schools? and (3)

How are the nation's schools responding to the introduction of the microcomputer--a technological innovation that may or may not revolutionize American education?

"Taking the Pulse" of American Education: What Are Students Doing and Learning in the Nation's Classrooms?

While data such as expenditures for public schooling provide important information on the nation's priorities and are useful to policy makers and others who must allocate resources, they do not provide an adequate measure of either the quality or success of our educational system. The criterion typically used to judge the effectiveness of our schools is students' achievement scores. Therefore, the achievement data collected by the National Assessment of Educational Progress (NAEP) are extremely important.

NCES and NAEP should continue to collect student performance data, and they should give increased attention to the need to improve the measurement of higher-order cognitive skills in reading, mathematics, and science. Although NAEP's measurement of higher-order skills far exceeds that of traditional standardized achievement tests such as the Iowa Test of Basic Skills, their measurement of higher-order skill is far from perfect. Both NCES and NAEP should give high priority to refinement and further development of test items that measure students' higher-level cognitive thinking in the major subject areas.

While students' achievement scores are an important measure of the condition of education, students' learning actually occurs in the nation's classrooms. To take the pulse of American education, we need to know what students are doing and learning in classrooms in the United States. The best metric to use in such an analysis is time.

Following the publication of the Nation at Risk report in 1983, many states responded to the recommendations by lengthening the school day; many school districts set minimal standards for the number of minutes that teachers must spend teaching each of the major subject areas during a given week. The impact of these new guidelines on what teachers and students are doing in classrooms has not been assessed. Moreover, the best data on time and content coverage were collected by the Beginning Teacher Evaluation Study in the mid-1970's (Fisher, Filby, Marliave, Cahen, Dishaw, Moore, & Berliner, 1978; Denham & Lieberman, 1980), and these data have not been updated since (Fisher & Berliner, 1985). Moreover, the original BTES data were collected on only selected grades (second and fifth grade) in a small number of schools in California.

Information is needed not only on the quantity of time allocated and spent in various activities in classrooms and schools, but also on the quality of the activity. For example, in preparing the recent report on the state of the art and practice in teaching reading in our schools, the Commission on Reading was unable to find information on the time that teachers are spending in phonics instruction in the early grades (Anderson, Hiebert, Scott, & Wilkinson, 1985). Information was also not available on the amount of time that students are spending in silent and oral reading in the elementary and middle school grades. According to the Commission Report, these measures might serve as indices of the effectiveness of the reading instruction that is occurring in our schools and would be highly related to student achievement in reading.

Similarly, although some researchers have documented that elementary students spend more than 50% of their time during reading and mathematics in seatwork activities (Fisher et al., 1978; Peterson & Fennema, 1985), little descriptive information exists on what students are working on during seatwork. In particular, we need to know the amount of time that students are spending in "busywork" which is unrelated to the academic subject matter, compared to the proportion of time that students are spending on specific academic content whether it be content that consists largely of drill and practice and requires lower-level cognitive thinking or whether it requires the students to engage in higher-level thinking. In their recent review of the research on school effectiveness, Good and Brophy (in press) pointed out that similar data are needed for homework that is assigned to students. A popular widespread belief today among the general public as well as educators is that students are not getting enough homework and need to spend more time on homework. However, no data exist on how much homework is assigned to elementary and secondary students each day, how much time students spend on homework, and the content of the homework that is assigned.

The above are examples of data that are needed on time spent by students in various instructional activities. In addition, data are needed on time allocated and student engaged time. (See, for example, Denham & Lieberman, 1980; Fisher & Berliner, 1985). The results of the BTES Study showed that while allocated time in reading and mathematics was significantly positively related to student achievement in reading and mathematics, student engaged time in reading and mathematics was a better predictor of achievement than allocated time.

How should allocated time, time spent, and student engaged time be assessed? Unfortunately, these data cannot be gathered by means of a survey or questionnaire administered once which appears to be the typical data collection technique used by NCES. Allocated time has been measured by having teachers log the amount of time that they schedule per week for a given content area. It is important that the content area be defined more specifically than simply reading, mathematics, or science so that the information will be useful to educational scholars and practitioners. Marliave, Fisher, & Filby (1977) reported that allocated time data from teacher logs "agreed at an acceptable level with the criterion of observational data" (p. 57). They noted further that the data were more reliable if they were recorded immediately after the event. Thus, perhaps teachers should not be asked to record the allocated time data for a whole month at a time, but rather for a week or several days at a time. In addition, although no researchers have investigated how many times during the year one must collect teacher logs on allocated time to get a generalizable estimate, one presumes that this information would need to be collected several times throughout the year.

Data on time spent and student engaged time must be collected through actual classroom observations. Although such observational data are costly to collect, they might provide a more valid measure of the condition of education than much of the survey data that has been collected by NCES in the past. Moreover, such information provides normative data on the quantity and quality of instructional practices that are occurring in various subject areas in our elementary and secondary schools as well as information on the quality and quantity of the content that students are purported to be learning. Such information would be useful for educational practitioners, policy makers, and researchers. In addition, researchers on effective teaching have found allocated time, time spent, and student engaged time to be significantly related to student achievement. Thus, these data may serve as potential indices of the quality of instruction that is occurring in American classrooms.

"Measuring the Blood Pressure" of the Nation's Teachers: What are the Stresses and Concerns of Teachers?

An upcoming crisis that may significantly affect the condition of education in the next decade is the severe shortage of qualified teachers. Data collected by NCES show that by 1988 the demand for teachers will far exceed the supply (National Center for Educational Statistics, 1984). The National Science Teachers Association estimates that 300,000 new mathematics and science teachers will be needed by 1995--more than the total number of mathematics and science teachers currently teaching (Darling-Hammond, 1984). Furthermore, reports have documented that the most talented teachers are leaving the profession. (See, for example, Schlechty & Vance, 1983; Darling-Hammond, 1984). These two factors may result in a teaching force that is considerably less qualified and competent than the present teaching force, which may have significant negative effects on the condition of education in the next decade. For example, studies of schools have shown that staff stability is an important measure of an effective school (New York State Department of Education, 1974; U.S. Department of Health, Education, and Welfare, 1978).

At the minimum, NCES should continue to collect the kind of data on supply and demand of teachers and turnover in the teaching work force that it has collected in the past. In addition, NCES should continue to collect data on teachers' salaries because salaries in the teaching profession have been identified as one of the salient factors related to retention of qualified teachers (Darling-Hammond, 1984; Schlechty & Vance, 1983).

Beyond these data, information is needed on what factors may be related to teachers staying in or leaving the teaching profession. Thus, NCES should collect data on: (1) the stresses and concerns of teachers in our nation's elementary and secondary schools; and (2) information on the professional working conditions of teachers in elementary and secondary schools.

Few large-scale survey studies have been done to document the concerns of teachers. However, most recent reports on schools (e.g., Boyer, 1983; Sizer, 1984) as well as the surveys that have been done (Darling-Hammond, 1984) suggest that teacher dissatisfaction and stress may be most related to professional working conditions. In addition to collecting large-scale survey data on teachers' concerns and stresses that may be related to teacher retention, NCES should also collect data on working conditions in the school.

In the Milwaukee School District, the largest urban school district in the state of Wisconsin, the two issues that are of greatest concern currently to teachers are: (a) class size; and (b) the amount of preparation time given to teachers. As a consequence, the Milwaukee Teachers Education Association is introducing legislation in the Wisconsin State Legislature to decrease class size of Milwaukee teachers and to increase their allotted preparation time. Inadequate preparation time was also one of three factors mentioned by the teachers in the Darling-Hammond (1984) study as causing the greatest teacher dissatisfaction.

NCES should continue to collect the kind of data that it has been collecting on teacher/pupil ratio and class size (NCES, 1985). NCES should also collect data on the amount of preparation and planning time given to teachers. This information might be collected through large-scale surveys of school districts and school principals. However, it is important to survey teachers about their preparation and planning time to check for validity and also for differences in perceptions of what constitutes planning and preparation time. In addition, data from teachers on how they spend their time during the day would provide some useful insights into the working conditions of teachers. For example, many studies have suggested that teachers are dissatisfied because they are overwhelmed with administrative duties and paperwork that detracts from the time they are able to spend in actual classroom teaching. (e.g., Boyer, 1983; Darling-Hammond, 1984).

"Assessing the Reflexes" of the American Educational System: How Are Schools Responding to the Microcomputer?

The recent influx of microcomputers into schools has stimulated widespread discussion and debate at all levels of our society. Indeed, the microcomputer has provided a focal point for contending educational philosophies and their attendant sets of priorities for allocation of funds and time within schools (Lepper, 1985). Enthusiasts believe that microcomputers will radically change education (Papert, 1980; Kleiman, 1984). Skeptics believe that the effects of the microcomputer are at best exaggerated and at worst will have negative consequences for education (Weizenbaum, 1976; Sloan, 1984; Brophy & Hannon, 1984).

There are several major reasons for gathering information on microcomputer usage and how schools are responding to the advent of the microcomputer. First, just as one aspect of an effective organization is its ability to respond to change (Chandler, 1962; Miller, 1978), one index of the condition and quality of our educational system may be the way in which schools are responding to the advent of an innovation such as the microcomputer. Second, the new age of advancing technology and global competition has radically changed our concept of "basic skills"--the skills necessary for a person's economic competence. Students will need to have basic skills in the use of microcomputers in order to function successfully as citizens in our society. In addition, the microcomputer is potentially a powerful tool for ensuring mastery of other basic skills, especially in the areas of reading, writing, and mathematics.

Third, information on how schools are using microcomputers provides an index of how our educational system is currently responding to issues of social equity. Although children from higher income families frequently have microcomputers at home, children from lower income homes rarely do. Further, schools serving more affluent communities frequently have greater parental pressure to acquire microcomputers, greater resources to buy them, and greater human resources for using them widely. If children from low income families are not given access to the new technologies in schools, they may fall even farther behind their affluent peers in their preparation for employment in a era of high technology. The same issues arise with respect to gender.

Considerable evidence exists that boys are much more likely than girls to become involved with microcomputers (Kisler, Sproull, & Eccles, 1983), especially at the more advanced levels (Hess & Miura, in press). Microcomputers are still a relatively scarce resource, especially in elementary schools. The educational community, policy makers, and the general public need to know how schools are responding to the challenge and opportunity to make effective and equitable use of this scarce resource.

Thus far, the only extensive national data collected on microcomputer usage in the schools is from a national survey conducted by (Becker, 1983a, 1983b, 1983c, 1984). These data were included in the 1984 Statistical Report of NCES. While survey data are useful on the number of computers owned by a school and the average amount of time per week that students spend on the computer, observational data are needed both to check the reliability of the survey data and also to provide a more complete picture of how microcomputers are actually being used in the schools. To address the issue of social equity, all data on microcomputer usage should be broken down by sex, race, and socioeconomic status (SES) of student.

The data collected on microcomputer usage in the schools should be similar to the kind of data described above on what students are doing and learning in the classroom. Such data might include: (1) allocated time on the microcomputer broken down by grade, sex, race, and SES of student; (b) actual time spent per week per student broken down by the same categories of student; and (c) student engaged time on the microcomputer per week. Allocated time data should be collected through teacher logs and logs of teachers who have responsibility for the microcomputer resource room. Allocated time and actual time spent should be clearly differentiated. For example, it is not clear whether in responding to Becker's national survey, schools were reporting weekly use data on microcomputers that reflected allocated time or actual time spent on the microcomputer. This question might be addressed by collecting observational data which could be used to check the reliability and validity of the survey data on time usage.

Data on time spent by each student on the microcomputer and student engaged time on the microcomputer should be collected through classroom observations and through observations in the microcomputer resource room. Observers should record the kind and content of the activity in which students are engaged while working on the microcomputer. Information is needed on: (a) how much time students are spending on actual academic activities compared to game-like activities; (b) whether the activity teaches computer literacy or a subject matter such as reading, mathematics, or writing; and (c) whether the microcomputer activity is a higher-level cognitive activity such as problem solving or a lower-level cognitive activity such as drill and practice.

Summary

In sum, we have argued that, in the next decade, NCES should collect data in three major new areas: (1) the quantity and quality of time that is allocated to various activities in the classroom, and the amount of time students are actually spending and engaged in such activities; (2) the concerns of teachers in our nation's schools and information on working conditions, including the amount of time that teachers are spending in various activities during the day; and (3) microcomputer usage in the schools, including allocated time, time spent, and student engaged time in activities broken down by content of activity and the sex, race, and socioeconomic status of the user.

Finally, in the next decade, NCES might improve the relevance, technical quality, and utility of their data as well as improve the cost effectiveness of their data collection efforts by collaborating with other large organizations, both in planning data collection and in gathering the data. One productive avenue for collaboration would be for NCES to work closely with several of the Educational Research and Development Centers that will soon be funded by the National Institute of Education (NIE). NIE has requested that each of these Centers reserve 10% of their budgets in 1987 through 1990 for collaboration with other national Centers (National Institute of Education, 1984).

If NCES were to focus its new data collection efforts on the issues emphasized in this paper, then NCES should explore collaborative relationships with the following Centers: (1) NIE Center on Teacher Quality and Effectiveness; (2) NIE Center on Student Testing, Evaluation, and Standards; (3) NIE Center on Effective Elementary Schools; and (4) NIE Center on Effective Secondary Schools. The NIE Center on Teacher Quality and Effectiveness is supposed to gather data from the teachers' perspective on working conditions and factors that affect teachers' decisions to stay in the profession. NCES might collaborate with the NIE Center on Student Testing, Evaluation, and Standards in the development and refinement of tests that measure higher-level cognitive skills in reading, mathematics, and science. Finally, NCES might work with the NIE Centers on Effective Elementary Schools and Effective Secondary Schools to collect time data and observational data on what students are actually doing and learning in elementary and secondary classrooms including data on usage of microcomputers.

References

Anderson, R. C., Hiebert, E. H., Scott, J. A., & Wilkinson, I. A. G. (1985). *Becoming a nation of readers: The Report of the Commission on Reading*. Washington, DC: The National Institute of Education, US Department of Education.

Becker, H. J. (1983a). *School uses of microcomputers: Reports from a national survey, No. 1*. Baltimore: Johns Hopkins University.

Becker, H. J. (1983b). *School uses of microcomputers: Reports from a national survey, No. 2*. Baltimore: Johns Hopkins University.

Becker, H. J. (1983c). *School uses of microcomputers: Reports from a national survey, No. 3*. Baltimore: Johns Hopkins University.

Becker, H. J. (1984). *School uses of microcomputers: Reports from a national survey, No. 4*. Baltimore: Johns Hopkins University.

Boyer, E. L. (1983). *High school*. New York: Harper & Row.

Brophy, J. & Hannon, P. (1984). *The future of microcomputers in the classroom*. (Occasional Paper No. 76). East Lansing, MI: Institute for Research on Teaching.

Chandler, A. D. (1962). *Strategy and structure*. New York: Academic Press.

Darling-Hammond, L. (1984). *Beyond the commission reports: The coming crisis in teaching*. Santa Monica, CA: The Rand Corporation.

Denham, C. & Lieberman, A. (1980). *Time to learn*. Washington, DC: National Institute of Education, US Department of Education.

Fisher, C. W., & Berliner, D. C. (1985). *Perspectives on instructional time*. New York: Longman.

Fisher, C. W., Filby, N. N., Marliave, R. S., Kahen, L. S., Dishaw, M. M., Moore, J. E., & Berliner, D. C. (1978). *Teaching behaviors academic learning time, and student achievement*. (Technical Report V-1, Final Report of Phase III-B Beginning Teacher Evaluation Study). San Francisco: Far West Laboratory for Educational Research and Development.

Hess, R. D., & Miura, I. T. (in press). Gender and socioeconomic differences in enrollment in computer camps and classes. *Sex Roles*.

Keisler, S., Sproull, L., & Eccles, J. S. (1983, March). Second-class citizens? *Psychology Today*, 41-48.

Kleiman, G. M. (1984). *Brave new schools: How computers can change education*. Reston, VA: Reston Publishing.

Lepper, M. R. (1985). Microcomputers in education: Motivational and social issues. *American Psychologists*, *40*, 1-18.

Marliave, R., Fisher, C. W., & Filby, N. N. (1977). *Alternative procedures for collecting instructional timed data: When can you ask the teacher and when must you observe for yourself*? Paper presented at the annual meeting of the American Educational Research Association, New York, New York.

Miller, J. B. (1978). *Living systems*. New York: McGraw-Hill.

National Center for Education Statistics. (1984). *The condition of education*. Washington, DC: U.S. Department of Education.

National Institute of Education. (1984). *Planning grants and institutional grants for educational research and development centers*. Washington, DC: The National Institute of education, U.S. Department of Education.

Fisher, C. W., Filby, N. N., Marliave, R. S., Kahen, L. S., Dishaw, M. M., Moore, J. E., & Berliner, D. C. (1978). Teaching behaviors academic learning time, and student achievement. (Technical Report V-1, Final Report of Phase III-B Beginning Teacher Evaluation Study). San Francisco: Far West Laboratory for Educational Research and Development.

Hess, R. D., & Miura, I. T. (in press). Gender and socioeconomic differences in enrollment in computer camps and classes. Sex Roles.

Keisler, S., Sproull, L., & Eccles, J. S. (1983, March). Second-class citizens? Psychology Today, 41-48.

Kleiman, G. M. (1984). Brave new schools: How computers can change education. Reston, VA: Reston Publishing.

Lepper, M. R. (1985). Microcomputers in education: Motivational and social issues. American Psychologists, 40, 1-18.

Marliave, R., Fisher, C. W., & Filby, N. N. (1977). Alternative procedures for collecting instructional timed data: When can you ask the teacher and when must you observe for yourself? Paper presented at the annual meeting of the American Educational Research Association, New York, New York.

Miller, J. B. (1978). Living systems. New York: McGraw-Hill.

National Center for Education Statistics. (1984). The condition of education. Washington, DC: U.S. Department of Education.

National Institute of Education. (1984). Planning grants and institutional grants for educational research and development centers. Washington, DC: The National Institute of education, U.S. Department of Education.

New York State Department of Education. (1974). School factors influencing reading achievement: A case study of two inner city schools. Albany, NY: Office of Education Performance Review. (ERIC Document Reproduction Service No. E.D. 089 211).

Papert, S. (1980). Mindstorms: Children, computers, and powerful ideas. New York: Basic Books.

Peterson, P. L., & Fennema, E. H. (1985a). Effective teaching, student engagement in classroom activities, and gender-related differences in learning mathematics. American Educational Research Journal, 22, 309-333.

Schlecnty, P. C., & Vance, V. S. (1983). Recruitment, selection, and retention: The shape of the teaching force. The Elementary School Journal, 83, 469-487.

Sizer, T. (1984). Horace's compromise: The dilemma of the American high school. Boston: Houghton Mifflin.

Sloan, D. (1984). On raising critical questions about the computer in education. Teachers College Record, 85, 539-547.

U. S. Department of Health, Education, and Welfare. (1978). Violent schools-safe schools: The faith school study report to the U. S. Congress. Vol. 1. Washington, DC: Government Printing Office.

Weizenbaum, J. (1976). Computer power and human reason: From judgment to calculation. San Francisco: W. H. Freeman.

New York State Department of Education. (1974). School factors influencing reading achievement: A case study of two inner city schools. Albany, NY: Office of Education Performance Review. (ERIC Document Reproduction Service No. E.D. 089 211).

Papert, S. (1980). Mindstorms: Children, computers, and powerful ideas. New York: Basic Books.

Peterson, P. L., & Fennema, E. H. (1985a). Effective teaching, student engagement in classroom activities, and gender-related differences in learning mathematics. American Educational Research Journal, 22, 309-333.

Schlechty, P. C., & Vance, V. S. (1983). Recruitment, selection, and retention: The shape of the teaching force. The Elementary School Journal, 83, 469-487.

Sizer, T. (1984). Horace's compromise: The dilemma of the American high school. Boston: Houghton Mifflin.

Sloan, D. (1984). On raising critical questions about the computer in education. Teachers College Record, 85, 539-547.

U. S. Department of Health, Education, and Welfare. (1978). Violent schools-safe schools: The faith school study report to the U. S. Congress. Vol. 1. Washington, DC: Government Printing Office.

Weizenbaum, J. (1976). Computer power and human reason: From judgment to calculation. San Francisco: W. H. Freeman.

ASSESSING NATIONAL DATA ON EDUCATION

Valena White Plisko
U.S. Department of Education
400 Maryland Avenue, S.W.
Washington, D.C. 20202

Alan Ginsburg
U.S. Department of Education
400 Maryland Avenue, S.W.
Washington, D.C. 20202

Stephen Chaikind
Decision Resources, Incorporated
1701 K Street, N.W.
Washington, D.C. 20006

August 1985

ASSESSING NATIONAL DATA ON EDUCATION

Gauging the condition of education in the United States necessarily involves some assessment of how well the gauges work. Are our information systems measuring key indicators? Do they provide accurate readings? Are they optimally linked? In this assessment, we apply these questions of coverage, quality, and linkages to the current collection of national statistics on education at the preprimary, elementary/secondary, and higher education levels. At each level, we discuss fundamental policy issues, summarize major data bases, and evaluate the potential of existing data to inform policy discussions. In the final section, we recommend ways to improve the collection of national statistics on education.

DATA ON PREPRIMARY EDUCATION

Concern about preprimary education in the United States has been stimulated by rising participation rates and a recent upturn in the population of preschool age. Heightening the concern is conflicting evidence over the importance of such education for later achievement. Results from the Perry Preschool Project (Berrueta-Clement et al., 1984) and earlier collaborative work (Lazar et al., 1974) indicate that preschool experience has sustained benefits for disadvantaged children. These results have been used to justify expansion of publicly supported preschool education. Critics, however, have argued that these findings are not applicable to most preschool programs.

The heightened interest in preschool education and disagreement over its effectiveness have generated a need for information on what education is currently being provided to preschoolers, what outcomes should be expected, and how preprimary programs complement learning in the home. A number of questions have been raised:

- Availability. What types of preschool programs are available? What types of children participate in various programs? To what extent is preschool education available to low-income families? How are programs supported? What is the mix of public and private funds? Is the current pool of preprimary programs adequate to meet the demand? Can it meet future demand?

- Standards. Under what standards do preprimary programs operate? Do standards differ by community, family background of participants, and sponsorship? What standards distinguish high-quality programs?

- Family-school interaction. How do preprimary programs complement learning in the home? How do families, in turn, reinforce preschool learning?

Only a half-dozen surveys collect nationally representative data relevant to these policy concerns. Table 1 shows the six data bases and the categories of information they cover. Responsibility for government-sponsored surveys is divided among various federal agencies. The Education Department surveys that gather some preschool information include the Common Core of Data (CCD) on public school systems and the High School and Beyond (HSB) longitudinal study of the National Center for Education Statistics, and the National Assessment of Educational Progress (NAEP) of the National Institute of Education. CCD provides statistics on kindergarten and prekindergarten enrollment in regular public school systems and full-time-equivalent teachers at this level. HSB and NAEP ask students to report retrospectively on whether they had been enrolled in preprimary programs.

The Bureau of the Census collects data annually on preprimary enrollment and occasionally on child-care arrangements of working mothers through the Current Population Survey (CPS). The Department of Health and Human Services (HHS) data collection on preprimary education is now limited to an annual Head Start survey, but in 1976-77 HHS also undertook a comprehensive study of day-care centers. Although the 1976-77 National Day Care Study was a one-time study, it may suggest a model for future data collections and is included among the data discussed here.

Data Coverage

Without knowledge of the current availability of preprimary education providers, it is difficult to assess how well enrollment demand is being accommodated now and is likely to be accommodated in the future. To answer these questions, policymakers need information on enrollments in various types of preschool programs, costs of services, and access of different population groups to preschool programs.

National data are available on participation in public and private nursery schools and kindergartens from the CPS. Despite the growing interest in private sector involvement in providing educational services, however, there is no information on the type of provider (whether nonprofit or profitmaking) and on sponsorship (whether church, employer, or community-sponsored).

The availability of services to different population groups may hinge on costs and the ability to pay for such programs but data are sparse on costs for preprimary programs. The National Day Care Study collected fairly extensive data on expenditures, revenue sources, fees, donated resources, and staff salaries, but these data applies to only one type of preprimary program, licensed day care, and the data are a decade old. In response to special requests from the Education Department, the CPS asked for information on tuition paid for private nursery schools and kindergartens in 1979 and 1982. But tuition data are not routinely collected in the CPS and can be tied to only a few program characteristics. The Head Start program annually estimates its average cost per child, but expenditures for Head Start cannot be generalized to other preschool arrangements. Moreover, the Head Start program serves only one in four of the eligible low-income children and provides no data on the total need for preschool care among low-income populations or on cost barriers to providing such care.

TABLE 1: VARIABLE CATEGORIES COVERED BY SELECTED PREPRIMARY EDUCATION DATA BASES

Data bases	Outcomes	Processes--curriculum, climate, standards	Resources	Student background
Education Department				
Common Core of Data (CCD)	--	--	◺	--
High School and Beyond (HSB)	◺	--	--	□
National Assessment of Educational Progress (NAEP)	□	--	--	◺
Other Federal Agencies				
Current Population Survey (CPS)	--	◺	--	□
Head Start annual survey	--	◺	□	□
National Day Care Study (1976-77)	□	□	□	□

□ = Complete data

◺ = Incomplete data

-- = No data

Questions of standards for preprimary education--and government involvement in setting standards--are being discussed in state legislatures across the nation. Informed debate requires knowledge of what constitutes adequate standards and how well existing services match these standards. Moreover, in defining an adequate level of service, it is important to know the extent to which preschool experience affects a child's short- and long-term progress in school. Questions to be addressed in establishing standards include these: What credentials should staff have? What represents an acceptable staff-to-child ratio? How much time should young children devote to structured learning activities?

Information on the quality of current programs is virtually nonexistent, lacking the most basic data on preschool processes and outcomes. The evidence from effective preschool programs points to the importance of the quality issue. The Perry Preschool Project, which yielded long-term benefits for young children from disadvantaged homes, stressed the "high quality" of the program, as measured by high staff-to-pupil ratios, extensive home visits, and a fully articulated curriculum.

The last national attempt to gather comprehensive information on preschool processes was the National Day Care Study in 1976-77. This study identified variables that appeared to improve children's gains in test scores, such as the specific child-related education of the caregiver, a finding that has implications for teacher certification.

The Head Start program's ongoing annual survey asks for information on staff credentials in early childhood education, but does not ask for any child performance measures beyond the number of children who drop out of the program, thereby missing an opportunity to link standards with outcomes.

Data from HSB and NAEP also could potentially tie participation in preschool with later performance, but retrospective reporting and the fact that nothing is known about the characteristics or quality of these programs limit the utility of these data.

"If studies of school achievement have shown one thing, it is the importance of the family," concludes Coleman in his latest study of schools (1982, p. 19). To assess the preschool experience, it is critical to discover the complementary nature of preschool-home relationships. For example, how much time do parents spend participating in learning activities with their children? What home learning activities are most beneficial? What are the net effects of preschool education in relation to the home environment?

Only one national data base provides any reasonable information on home activities for children. The University of Michigan Time-Use Study gathered information in 1975 and 1981 on parent-child interaction through home diaries, but this study included too few young children to focus on the preschool years and did not collect performance data for this age group.

Indeed, no national surveys have looked at the home environment and the preschool environment together to suggest how learning may be reinforced or confounded by different settings. Some researchers have posited that the "hidden curriculum" provided in middle-class homes exerts so strong an influence on learning that instruction offered in a formal preprimary setting yields no net learning benefits for middle-class youngsters. Current national data bases provide no information with which to address this subject.

Data Quality

Data availability is obviously the largest problem at the preprimary level, but accuracy is also a concern. When the various data bases do overlap, they may not agree. For example, public preprimary enrollment reported through CCD shows some half-million children fewer than the 3.5 million estimated from the CPS household survey.

An examination of the CCD suggests that some local school districts and state education agencies have difficulty reporting accurate preprimary enrollment and staffing information, particularly in distinguishing between headcount and full-time-equivalent enrollments. Yet household respondents in the CPS may be equally inaccurate in reporting preprimary enrollment. Although the CPS distinguishes between home day care and formal preprimary programs, it leaves the classification of center-based day care to the household respondent. Thus, the rather thorny problem of defining center care as instructional or custodial is left to the respondent and presumably this definition could vary from one respondent to the next.

Data Linkages

A major hindrance to progress in developing data on preprimary education is the fact that no single agency has a clear mandate to collect the information. The Department of Education is just beginning to formulate proposals to address the area of early childhood education. The largest federal education program at the preprimary level, Head Start, is administered outside the Department of Education by HHS. In recent years, HHS has all but closed down its statistical collection activity in preprimary education, except for annual reporting on Head Start. The comprehensive but costly ($7 million) National Day Care Study in the late 1970s required a commitment of resources that are no longer available. With no agency yet assuming the lead in this area, data collection efforts have been sporadic, piecemeal, and uncoordinated.

DATA ON ELEMENTARY AND SECONDARY EDUCATION

Concerns with improving education dominate issues in elementary and secondary education. Various reports have called attention to performance declines over the recent past and the need to upgrade American education, spurring many states to pass costly reforms. Policymakers, educators, and the public need to be able to assess whether these reforms are working and whether they are getting their money's worth. In addition, the drive toward excellence has heightened the concern about providing equal opportunity to all students. It is generally recognized that if the Nation's schools are to improve, everyone should be encouraged to share in this improvement. Current policy debates revolve around these questions:

- Educational improvement. Are educational reforms actually improving student outcomes?
- Resource effectiveness. Are resources being applied to promote cost-effective educational programs?
- Equity. Are the needs of disadvantaged and other special population groups being met?

Although a number of national data bases describe different aspects of elementary and secondary education, three that have already been mentioned stand out as the major ongoing sources of national information, as shown in Table 2. The CCD survey of all state educational agencies and local school districts, is the primary source of information on schooling inputs--enrollment, finances, and staffing. NAEP focuses on measuring educational outcomes. The HSB longitudinal survey is the only one with extensive information on family background, school process variables, and student outcomes. Specialized information on elementary and secondary education is available from various other NCES surveys: the Private School Survey; the Survey of Teacher Demand and Shortages; and the Library/Media Center Survey.

With respect to particular populations of special concern to the federal government, the Office for Civil Rights surveys districts and schools for information on programs, disciplinary actions, and graduating class composition. Individual federal education programs, such as those serving the handicapped or the limited-English proficient, conduct their own surveys of participants and coverage.

Other federal agencies and private organizations also provide elementary and secondary education data. The CPS, mentioned earlier, supplies annual data on enrollment, public and private, and educational attainment. The Justice Department collects information on school discipline. HHS funds "Monitoring the Future," an annual survey of high school seniors which focuses on student attitudes and drug abuse. The National Science Foundation (NSF) conducts studies on math and science education. The Labor Department's National Longitudinal Survey--Youth Cohort (NLS--Youth) contains extensive information on family, schooling, and work history. A private consortium, the International Association for the Evaluation of Educational Achievement (IEA), periodically assesses comparative student performance in various nations. The National Education Association (NEA) collects information from its state coordinators on enrollment, staff, and salaries.

Data Coverage

As a result of the clear demand for school improvement and the expenditure of billions of dollars on educational reforms, information is critically needed on whether these reforms are improving outcomes and which reforms seem most effective. Ultimately, data should indicate whether student performance, as measured by educational achievement and attainment, has improved and for

TABLE 2: VARIABLE CATEGORIES COVERED BY SELECTED ELEMENTARY AND SECONDARY EDUCATION DATA BASES

Data bases	Outcomes	Processes--curriculum, climate, standards	Resources	Student backgrour
Education Department				
Common Core of Data (CCD)	--	--	◺	--
National Assessment of Educational Progress (NAEP)	□	◺	◺	□
High School and Beyond (HSB)	□	◺	◺	□
Private School Survey	--	--	◺	◺
Teacher Supply and Demand Survey	--	--	◺	--
Library/Media Center Survey	--	--	◺	--
Civil rights survey	◺	◺	--	◺
Special education program surveys	--	◺	◺	◺
Language minority surveys	--	--	--	◺
Other Federal Agencies				
Current Population Survey (CPS)	◺	--	--	□
Justice Department surveys	--	--	--	□
National Science Foundation (NSF) surveys	--	◺	◺	--
National Longitudinal Survey--Youth Cohort (NLS--Youth) (Labor Department)	□	◺	--	□
Monitoring the Future--Drug Abuse (Health and Human Services)	--	◺	--	□
Private Organizations				
International Association for Evaluation of Educational Achievement (IEA)	□	□	--	--
National Education Association (NEA)	--	--	□	--

□ = Complete

◺ = Incomplete data

-- = No data

whom. This information should be available nationally, but also disaggregated at least down to the state level, since the states are primarily responsible for education.

Most of the national data bases listed in Table 2 provide little help in answering these questions. Few provide performance information. The extensive data collected on resources can report little about the quality of inputs, such as the quality of teachers. A few data bases contain some information on schooling processes, but the data often apply to highly specific areas such as the provision of special education for the handicapped.

HSB presents reasonably good information on school outcomes and processes, but its performance test is limited to only a few items. The information pertains only to high school processes and tells us little about experiences before high school. The national data are of little help to particular states working to evaluate their systems.

NAEP offers some advantages over HSB in assessing reforms. NAEP's tests are much more extensive than HSB's and are not limited to high school students. NAEP also surveys the schools more often than HSB. Furthermore, although NAEP initially focused on outcome data, it has expanded its information on classroom processes and school and teacher characteristics, and it contains some home background information.

But NAEP also has disadvantages. First, NAEP is subject specific, that is, in a given year, NAEP focuses primarily on a few learning areas instead of providing a comprehensive picture of what is happening within the whole school. Second, NAEP consists of repeated cross sections and cannot be as informative as HSB with respect to measuring transitions over time, such as dropping out of school. Furthermore, like HSB, NAEP cannot provide state- or district-level comparisons. The current NAEP design does allow states to expand the sample to provide representative scores, but to date only three states have done so. Thus, although states carry the primary responsibility for education, the nation has no way of accurately assessing how well state governments are carrying out these responsibilities.

Other data bases provide additional information on reform, but all are limited in important respects. IEA compares achievement in a number of countries, but the data are highly aggregated and the long intervals separating reports (sometimes 15 years) makes IEA unsuitable as a gauge of reform effects.

On the question of resource effectiveness, research findings have shown that overall expenditures and student performance are only weakly associated. Hence, to examine the quality of resources and the ways they translate into school processes, statistics must be collected below the level of broad expenditure components.

The CCD annually collects revenue, expenditure, and debt-financing statistics on some 16,000 school districts. Some of these data are important for assessing efficiency issues, such as breakouts of administration and instructional components of expenditures, but many of the items hardly warrant reporting each year. On the other hand, despite the concern over the quality of teachers--the most critical resource, CCD no longer collects information on what each state pays its teachers.

Because administrative records on inputs are readily available, data are often collected without sufficient regard to priorities. Library statistics, for example, are able to provide data on a full range of variables. Although libraries are essential for education, something is wrong when NCES can report periodicals' costs but not teachers' salaries.

Measuring how well groups with special needs, such as language and racial minorities, are faring as a result of reform efforts requires performance information for these groups and some understanding of the relationship between education in school and conditions outside school. Data on special-need populations and their performance are available from program data, HSB, and NAEP, but the data are inaccurate. And we lack the information to understand why children with special needs fall so far behind in school, despite evidence of progress in the earliest grades. If we are to measure the extent to which students fall behind early, we need detailed information on the school and outside environment in the formative pre-high-school years, data that are almost nonexistent.

Data Quality

The quality of data on elementary and secondary education varies greatly. Performance, as indicated by standardized tests, is one of the better-measured variables. NAEP, in particular, has devoted considerable effort to improving outcome measurements and has introduced a new scaling system that may permit comparisons across grades.

Estimates of high school dropouts are less satisfactory. In the drive to promote educational excellence, the high school dropout issue was initially swept aside. This appears to have been a serious mistake; evidence from several sources shows that the problem is severe. However, the reported high school dropout rates are error prone. Estimates from alternate sources differ on the extent of the problem and, in the absence of agreement on a baseline number, it is hard to gauge whether the dropout problem is improving or getting worse.

A dropout rate can be calculated from the CCD by comparing the residual of the number of high school graduates with the number of students in the 9th grade 4 years previously; this approach yields a dropout rate of 27 percent. In contrast, the CPS, which uses household interviews for information on educational attainment, reports a dropout rate of only 16 percent. The CPS does not, however, report how schooling was completed, whether through graduation from a regular high school program or other means, such as earning an equivalency certificate or external degree. The dropout rate can also be calculated from the HSB longitudinal study, but this rate is generally acknowledged to represent an undercount because it misses students who leave school before the end of their sophomore year.

School process variables pose similar problems; even the simplest are sometimes hard to measure. HSB, to its credit, compared student self-reports of courses and grades with coursework transcripts and found substantial misreporting. Unfortunately, quality control studies have not been carried out for other types of process data, such as information on school climate and order. Students' responses need to be validated against reports from teachers and administrators on these critical variables.

Data on the use of resources are so unreliable that assessing efficiency in American schools is quite problematic. Large, implausible year-to-year variations in per pupil expenditures appear for some states. Pupil-to-staff ratios calculated from state reports also are unreliable; for example, the highest and lowest state pupil-to-all-staff ratios differed by 140% on this measure in 1983. Either the reported numbers are faulty, or some states have remarkable efficiency advantages over others.

On the question of whether special populations are being well served by American education, the data are also suspect. A prerequisite for addressing this issue is agreement on the target populations; numbers for most of the federal target groups are at best questionable. For example:

- In 1984, special educational students as a proportion of State enrollment ranged from a low of 5% to a high of 13% (U.S. Department of Education, Office of Special Education and Rehabilitative Services, 1985). There is no physiological explanation that could account for these report differences exceeding 100% in the prevalence of handicapping conditions.

- The Office of Bilingual Education and Minority Languages Affairs estimates that about 3.6 million language-minority children need bilingual services, but a recent study based on Census Bureau data concludes that two-thirds of the children included in this number use English as their usual or sole language, and thus could hardly be classified as candidates for instruction in a language other than English.

Finally, there is the difficulty of obtaining educationally relevant information on households, such as family income and parental reinforcement of education. This information appears to be far more accurately reported through home questionnaires rather than from student responses. HSB validity studies show that students may seriously misjudge family background characteristics; comparisons between students' and parents' reports of parental occupation, family income, and mother's work, for example, yielded validity coefficients of only about .5 (U.S. Department of Education, National Center for Education Statistics, 1984).

Statistical gathering by the federal government often involves inordinate time lapses between data collection and dissemination. Indeed, private organizations routinely publish similar data that are both more comprehensive and more timely. For example, the CCD provides the location and enrollments of all elementary and secondary schools, but the most recent school year for which this information is available is 1982-83. In contrast, private firms have compiled similar files that provide much greater information on each school site for the 1984-85 school year. As a second example, the NEA regularly publishes state-level financial and staffing data for the preceding and current school years. These data are generally more complete than those the Education Department publishes and at least 1 to 2 years more current.

Data Linkages

The collage of elementary and secondary data bases provides cogent examples of the advantages of integrating performance data with other kinds of information. Two of the most successful information activities in the elementary and secondary school area, in terms of use and attention to findings, are HSB and the Education Department's "wall chart" which graphically profiles state-by-state comparisons on college entrance test scores and resource variables. By contrast, detailed financial statistics reported in isolation in the CCD or extensive library statistics tell little about school quality and, hence, are not widely used.

The data on teachers also points to the need to consider data collection as a cohesive whole. Five separate questionnaires from the Department of Education have recently sought information from the states about teachers: CCD, HSB, NAEP, Survey of Teacher Demand and Shortage, and a new Public School Survey. Yet after all these surveys, the Department still does not know how much each state pays its teachers.

DATA ON HIGHER EDUCATION

Education Secretary William J. Bennett has called for a rethinking of higher education--both its funding and its functions. Questions have been raised concerning access, cost, quality, and the role of higher education in the United States today. Because student aid accounts for nearly half of all the Department of Education's spending, the government needs to know if the aid is being properly targeted and wisely spent. Institutional efficiency is also being questioned in relation to rapidly rising tuitions and purported decreases in faculty teaching loads. Indications of declining performance and unbalanced curriculums, moreover, call into question the very quality of higher education. The trend toward greater vocationalism in college studies, along with increased corporate training, stimulates concern that higher education is not fulfilling its mission.

Data sources presented in Table 3 are considered in this section as they pertain to these issues:

- Student aid. Who receives federal student aid and how can aid be refocused to provide greater access to the most disadvantaged students?

- Efficiency. Why have tuition charges increased faster than the inflation rate, and how can escalating costs be contained?

- Value. What is the quality of the educational experience and of college graduates? What is being learned?

- Role. What is the role of higher education, particularly in relation to occupational and corporate training?

Information on higher education is obtained by a diverse group of government and private organizations. At the federal level, the Higher Education General Information Survey (HEGIS) of NCES collects information from all institutions of higher education on enrollment, institutional control and finance, degrees conferred, and faculty employment and salaries. Other information is obtained by a series of surveys also sponsored by NCES; these include the National Longitudinal Study of the High School Class of 1972 (NLS), HSB, and the Survey of Recent College Graduates. In an October supplement the CPS collects data on college enrollment by control of institution (public or private), and family demographic and social characteristics. NSF surveys gather data on scientific and engineering education. The Labor Department's longitudinal study (NLS--Youth) tracks the work and education experience of young adults.

Sometimes specific information can be obtained only from sources outside of the government. The College Board's Annual Survey of Colleges, for example, is a comprehensive survey that queries institutions on admissions policies, student charges, standards, and fields of study. The American Council on Education (ACE) conducts approximately six Higher Education Panel (HEP) surveys each year on topics of special interest. The National Institute on Independent Colleges and Universities (NIICU) surveys student aid in private colleges; the National Association of State Colleges and Universities collects an analogous public college survey. The Cooperative Institutional Research Program (CIRP) surveys first-time freshmen, while the National Association of College and University Business Officers (NACUBO) quantifies endowment levels. As further examples, the American Association of University Professors (AAUP) and the Council for Financial Aid to Education (CFAE) report faculty salaries and voluntary support to higher education, respectively.

Data Coverage

The Reagan administration's 1986 budget recently proposed modifications in the federal student aid programs to better target benefits to the neediest students. Documenting the trends in aid recipients, though, has proved difficult and controversial. The problem is not that the information is insufficient, but that different sets of data are disjointed and cannot describe the total student aid package in relation to students' financial need. Once again, the problem is linkage rather than coverage.

The efficiency of the higher education production process is another issue. In recent years, college tuitions have risen faster than the inflation rate. Has efficiency in the production of higher education changed? Information to address this question would certainly include data on credit hours taught per faculty member, time divided between research and teaching, and class size and unit costs of different course offerings.

HEGIS, the principal federal data base to address these questions, collects data on number of faculty members and salaries, enrollment, revenue, and expenditures. Yet HEGIS data do not help us answer questions of efficiency. Faculty salaries are not reported according to the amount of time spent on research, as opposed to teaching. And information is not provided on credit hours taught or on class size from which unit costs by course offering could be calculated.

HEGIS fiscal data are aggregated at the institutional level, a practice that precludes reporting resource allocations among academic departments and levels. The extent to which costs, for instance, in the sciences are subsidized by humanities departments, or graduate programs by undergraduates, is not available. In addition, although it is generally assumed that the expansion of programs and course offerings has been costly and inefficient, no figures are available with which to measure the expense of offering a multiplicity of courses.

Alternative sources for certain information about higher education inputs are available, but little is known about the efficiency with which resources are allocated. The AAUP publishes annual salary statistics for college faculty by rank but reports nothing on teaching loads. It has been suggested (Noah, 1985) that teaching loads have been cut substantially over the past four decades, yet this claim cannot be corroborated because no national survey divides faculty time among teaching, research, and other endeavors.

There have also been recent expressions of concern about the value and diversity of the college course of study. "Although more than 50 percent of America's high school graduates continue their education at American colleges and universities," writes Secretary Bennett (1984), "few of them can be said to receive there an adequate education in the culture and civilization of which they are members." Information on course enrollments by subject area and on the extent of remediation would be helpful. Test scores would indicate what students actually learned in courses. The apparent growth in remediation would suggest a "dumbing down" of coursework in colleges. Although detailed information is available about the major fields of graduates, little is known about the courses graduates took outside their fields. For example, HEGIS reports the number of men and women receiving degrees in home management, but not the number of courses taken by students who major or do not major in the humanities. ACE surveys indicate total credit hours in selected science and humanities programs, but the data do not permit translation of these credit hours to individuals.

TABLE 3: VARIABLE CATEGORIES COVERED BY SELECTED HIGHER EDUCATION DATA BASES

Data bases	Outcomes	Processes	Resources	Student background
Education Department				
Higher Education General Information Survey (HEGIS)	◺	--	□	--
National Longitudinal Study of the High School Class of 1972 (NLS)	◺	--	◺	--
High School and Beyond (HSB)	◺	--	--	□
Survey of Recent College Graduates	◺	--	--	◺
Other Federal Agencies				
Current Population Survey (CPS)	◺	--	--	◺
National Science Foundation (NSF): Surveys of Science and Engineering Expenditures, Federal Support, Personnel, and Graduates	◺	--	◺	--
National Longitudinal Survey--Youth Cohort (NLS--Youth) (Labor Department)	□	◺	--	□
Private Organizations				
National Institute on Independent Colleges and Universities (NIICU): Survey of Student Aid Recipients	--	--	◺	□
National Association of State Colleges and Universities (NASCU): Survey of Student Aid Recipients	--	--	◺	□
College Board: Annual Survey of Institutions	--	□	◺	--
Cooperative Institutional Research Program (CIRP): American Freshman Norms	--	--	--	◺
American Council on Education (ACE): Higher Education Panel (HEP)	--	◺	◺	--
National Association of College and University Business Officers (NACUBO): Comparative Performance Study and Investment Questionnaires	--	--	◺	--
American Association of University Professors (AAUP)	--	--	◺	--
Council for Financial Aid to Education (CFAE)	--	--	◺	--

□ = Complete data ◺ = Incomplete data -- = Blank

Data that would quantify the returns to a college education in terms of increased knowledge also are lacking. No uniform test, such as the Graduate Record Examination, is given to a representative sample of graduates to test their general knowledge. One study of changes over the past two decades found significant declines in graduate admissions test scores in most fields (Adelman, 1984), but these results were not drawn from statistically representative sample of all graduates, because only those intending to go to graduate school took the exams.

Of concern is the quality of graduates produced by various programs and in different fields, particularly graduates of teacher education programs. Beyond the NLS study of members of the 1972 high school class who went into teaching, little definitive is known about the quality of graduates prepared to teach. The Survey of Recent College Graduates asked for grade-point averages, but this measure is a poor indicator of quality and no substitute for college transcript data. College transcripts would give not only grades but actual courses that would suggest the extent to which teaching graduates are fully prepared in their subject matter or were enrolled in watered-down courses.

Concern has also arisen over the mission of higher education in relation to other providers of postsecondary training. Some observers charge that higher education, particularly in community colleges, has become indistinguishable from vocational training. Conversely, corporate training is being provided in areas traditionally reserved for universities and colleges. These developments suggest the need for information about the extent of the overlap and the relationship between training provided by businesses and postsecondary institutions. For that matter, the academic-occupational mix in college programs remains an important unknown, critical to assessing higher education's role in relation to the role of other postsecondary providers.

Data Quality

Although many data sets provide information about the specific policy issues raised here, multiple sources often result in conflicting and incomplete overall higher education policy evaluations. Both the HEGIS and CPS, for example, report enrollment, but tabulations differ even for this most basic statistic. The two surveys basically agree on total enrollment but show discrepancies by full-time/part-time status and 4-year/2-year disaggregations and wider differences by graduate/undergraduate breakdowns. For example, CPS estimates graduate school enrollment to be one-third larger than the figure reported through HEGIS, although in other categories CPS finds fewer students than HEGIS.

Although HEGIS remains the federal government's primary instrument for monitoring higher education, the periodicity of many of these surveys has been interrupted in recent years. Because of processing delays and technical difficulty, NCES has abandoned attempts to release the 1982-83 data on faculty salaries and some of the financial and degrees data. Even in years without abnormal delays, HEGIS data were not available to analysts until approximately two years after the surveys were taken.

Data Linkages

A major problem associated with higher education data is that information from multiple sources cannot be combined to form a complete picture. The relationship between student aid and financial need provides an example of the lack of linkage among data sets and the difficulty this causes to analysts.

Any evaluation of student aid programs requires an understanding of the income distribution of students' families and the types and amounts of aid these students receive. These are important policy considerations, especially given the debate over proposed federal funding reductions.

The fundamental problem with student aid information is that it is collected for administrative purposes by those responsible for each program. Hence, coverage may not be a problem, but organization is. Because the information is generally not consolidated into files that cover all student aid programs, it is impossible using federal data bases, to determine how the need for student aid is distributed and how federal aid combines with other revenue sources to meet this need.

Two private sources of student aid information, however, cut across programs and could provide the needed data. Identical questionnaires concerning student aid are sent to both public and private institutions for the National Association of State Colleges and Universities (NASCU) and the National Institute on Independent Colleges and Universities (NIICU), respectively. These surveys could, if merged, provide the information needed for policy analysis and discussion. The Education Department can obtain the NASCU public college data, but NIICU will not make available the private college data tapes.

The problem, again, is that much of the information useful for analyzing the higher education process is piecemeal, serving only the specific needs of the originating agency. Some national coordination for these blocks of data would vastly improve the information base.

RECOMMENDATIONS FOR IMPROVING NATIONAL DATA

The tremendous national interest in educational improvement provides the impetus for reform of our national base of educational statistics, but the task will not be easily accomplished. Inertia and special interests will work against developing a coherent data collection strategy. Policymakers may be unwilling to wait for real improvements.

Given these problems, identifying a clear agenda becomes all important for improving national statistics. Extensive analysis is needed before a complete set of reforms can be specified. Analysts must carefully investigate the design options for data collection before final decisions on implementation are made. Resource and political constraints also will impose trade-offs among options. Our purpose in setting forth this agenda is to identify the most important directions for reform. We have divided our recommendations along the three types of evaluation criteria used throughout this paper: data coverage, data quality, and data linkages.

Data Coverage

The federal government needs to ask the important questions and collect data accordingly. It already collects a considerable amount of educationally related data, but some of the data now collected may have relatively little policy value, whereas other data not now gathered may have high informational value.

Recommendation 1: Eliminate low priority data items.

The continued collection of data with little policy worth cannot be justified. A comprehensive reconsideration of survey instruments requires a item-by-item analysis, but various candidates for paring are evident. They include the following:

- o Annual HEGIS collection of degrees awarded by sex, for six-digit degree codes at the subbaccalaureate level that include facts such as the number of female degree recipients in fashion merchandising or jewelry marketing;
- o Library statistics on book binding costs and number of audiovisual materials; and
- o Annual CCD collection of school district finances detailing expenditure and revenue statistics for all 16,000 school districts.

Recommendation 2: Identify and fill information gaps on a priority basis in areas of major policy interest.

Some gaps in the data can be filled by adding items to existing surveys without incurring substantial costs, and these efforts should proceed. Examples include:

- o Adding questions on preprimary education to the CPS to differentiate provider types and extending coverage of this item to estimate the number of children in home day care;
- o Requesting teacher salary data by state as a regular part of the CCD;
- o Including on the CPS a question on how high school was completed, whether through graduating from a regular 4-year program or by earning an equivalency certificate or external degree;
- o Adding questions to HEGIS about number of faculty hours spent teaching as a rough measure of instructional load; and
- o Adding a college transcript collection to the Survey of Recent College Graduates.

Supplementing existing surveys will not, by itself, produce much of the new data needed to resolve many of the most glaring deficiencies in policy information needs. These deficiencies include the following:

- Preprimary education. Little is known about the organization and support for preprimary education, the nature or quality of these programs, and the relationship between preschool and home learning.
- Elementary school. No recent study describes how course content relates to outcomes. Of special interest at this level would be information on the use and effectiveness of alternative teaching approaches to development of reading and mathematics skills. NAEP may begin to yield information on coursework in the 4th grade, but additional data are needed to describe how learning at this level relates to learning before and after this grade.
- Junior high school. Two areas deserve special attention: First, the adolescent years are ones in which behavioral problems of students become serious. Many youths may drop out of school at this level, yet national data about attrition in junior high is lacking. Information on how discipline and dropout problems develop may require longitudinal data collection at the junior high level. Second, deficiencies in higher-order thinking skills surface in adolescence, and these deficiencies need to be related to course-taking and schooling processes.
- Higher education student aid. The primary problem is the absence of a file for each student that shows aid from all sources in relation to need. This consolidated record is essential for analyzing the effects of student aid reforms.
- College learning. Except for the number of degrees granted and the unrepresentative graduate admission exam scores, there is little or no information on the learning that takes place in college or even on the distribution of coursework.

This rather long list unfortunately reflects the sorry condition of current national statistics. In view of financial and staff limitations, priorities must be established. Attention must be paid to state-of-the-art problems in developing designs that yield reasonably cost-effective information payoffs. Examples of the difficulties involved in conducting surveys include the following:

- o Preprimary education occurs in myriad settings, the most important of which may be the home. The universe of settings is poorly defined; master lists to draw samples often do not exist. Distinctions between nursery, kindergarten, preschool, and day care may be only nominal. Studying classroom processes at the preprimary level may require the use of detailed and costly observational techniques.

- o Previous attempts to measure schooling processes at the elementary level have yielded little useful information. The Sustaining Effects Study, a $20 million longitudinal study in the mid-1970s, thoroughly examined elementary schools, yet told very little about effective elementary schools and classrooms. For that matter, it even failed to achieve its primary objective of settling the debate over the long-term effectiveness of compensatory education programs.

- o Colleges, in particular, and postsecondary institutions, in general, are extremely diverse, and it is not at all clear how to measure outcomes for these institutions.

In view of these methodological and definitional considerations and differing policy priorities, experts will disagree over subjects to pursue. Among this list, the two activities we recommend for highest initial priority are these: (1) the junior high school study, because schooling in these grades may be the key to rejuvenating the high school and because so little is known about processes at this level; and (2) the student aid study, because this issue is of extreme immediate importance. In other areas, conceptual and measurement efforts are clearly needed and should begin at once. For example, efforts to separate educational from custodial services at the preprimary level and to develop of appropriate measures of learning at the college level should be pursued.

Data Quality

Inaccuracies and inconsistencies in Education Department statistics pose serious problems. Once data are published, they become accepted as fact, regardless of the caveats that may accompany their publication. In some instances, the reader is not even warned of data weaknesses. These concerns over data quality are not new ones, but seemingly they arise, year after year, with little improvement. The Education Department should immediately begin to address the problems of improving the adequacy of its data.

Recommendation 3: Establish an office of quality control.

No such office now exists within the Education Department, nor has any office promulgated a set of standards to ensure adequate data quality. An office that has this function as its primary responsibility is essential to the job of improving data quality. This office should be independent of the data-collecting offices and should oversee data collection of both general purpose statistics and program data.

Recommendation 4: Give highest priority to improving the quality of elementary and secondary education data.

Although all information must be subject to quality control reviews, the most glaring accuracy problems relate to elementary and secondary education data. The states must be held accountable for providing the Department of Education with consistent and accurate data.

The failure of certain data to pass muster when subjected to external validity checks must be immediately addressed. In particular, differences between NCES and Census Bureau data on estimated dropout rates and private school attendance can no longer be ignored simply because of organizational divisions at the federal level. The two organizations should immediately establish a joint review group to assess the reasons for inconsistencies in the data each collect.

Recommendation 5: Modernize procedures for data collection.

NCES has had problems of providing timely data, especially with some of its larger surveys such as CCD and HEGIS. Several types of reforms could improve the currency of information. Using the telephone instead of a mailed questionnaire would speed collection of school universe data. When there are a limited number of respondents, such as 50 states, computer networks seem to be a sensible approach. Once data are collected, they should not languish. Analysis contracts should be built into the overall data collection effort to expedite reporting of the data.

Data Linkages

The number of distinct data collection instruments related to education is quite large. The approximately 35 data bases reviewed in this study are only a partial listing. These activities lack cohesiveness. Although a master plan relating all data collection efforts would be quite unwieldy, far too little attention is now paid to the advantages that could be derived from coordinating or combining related efforts.

No one reason explains why coordination of national data-gathering activities has not progressed. Sometimes the reasons are historic. New data collections were planned for specific purposes and the planners paid inadequate consideration to whether these purposes could be better met through existing activities. Moreover, data needs may be defined so narrowly that the benefits of an integrated data set are not perceived. At other times, the problem is bureaucratic. It is unnatural to expect one statistical division or agency to transfer its responsibility to another. Whatever the reason, the limited federal funds available for education statistics make coordination of data activities imperative.

Steps the Education Department should consider to strengthen linkages among its statistical activities include the following:

Recommendation 6: Investigate the feasibility of linking the most costly data collection systems in the Department of Education.

Longitudinal data and performance assessments are acknowledged to be high-budget items. If we can detach ourselves from the ways data systems are currently organized and administered, we may see some alternatives for collecting information. Take, for example, the NAEP. Many would say that NAEP has had limited usefulness because it could not be linked directly to school policies and practices. Prior to the 1983-84 assessment, performance measures could be tied to only a few student characteristics and to no curriculum or process variables. The recent release of preliminary NAEP data suggests that the inclusion of fairly comprehensive student and teacher questionnaires will prove most valuable.

Other tie-ins to major outcome assessment activities should be considered. In particular, the question should be raised as to whether the distinctions between repeated cross-sectional studies, such as NAEP, and longitudinal studies, such as HSB, are real or artificial. For example, attaching a small longitudinal component to the NAEP 7th grade sample might provide a measure of the extent of attrition at this early level. The relationship among longitudinal surveys also should be considered. such as, for instance, how NCES longitudinal surveys could be coordinated with the Labor Department's NLS--Youth survey which also obtains longitudinal cohort data. Both HSB and Labor's Youth Cohort have transcript studies that, coincidentally, are being directed by the same contractor.

Recommendation 7: Initiate informal discussions with representatives of Departments of Education, Health and Human Services, and Labor, and the Bureau of the Census to coordinate data collections.

Data collectors in the private sector also should be brought in to the discussions.

Recommendation 8: Consider, when developing questionnaire items, those variables that have been shown to be most important to educational outcomes.

Statisticians and survey monitors should draw upon recent school effectiveness studies to help frame survey instruments.

Recommendation 9: Establish an education data bank to improve survey consistency across data bases and over time.

This data bank would pool items related to education within and outside government and might include items used in state and local surveys and case studies.

Recommendation 10: Keep expectations high.

We might take a cue from recent research on HSB which shows the power of positive thinking and persistence on student performance. Despite limited funds, national data collections can be improved. Because funds are limited, greater efficiency it is important to promote in our information systems. In addition, data producers should call attention to and take credit for the policy payoffs from study findings.

In addressing 1985 college graduates, Secretary Bennett offered some sound advice: "It is practical optimism that I recommend." As we consider reforming national statistical collections, we must think practically and optimistically about the task ahead.

REFERENCES

Adelman, C. (1984). The standardized test scores of college graduates, 1964-1982. Prepared for the Study Group on the Conditions of Excellence in American Higher Education.

Bennett, W. J. (1984). To reclaim a legacy: A report on the humanities in higher education. National Endowment of the Humanities, Washington, DC: U.S. Government Printing Office.

Bennett, W. J. (1985). Commencement address. Released by the U.S. Department of Education, Office of Legislation and Public Affairs.

Berrueta-Clement, J. R., Schweinhart, L. J., Barnett, W. S., Epstein, A. S., & Weikart, D. P. (1984). Changed lives: The effects of the Perry Preschool Program on youths through age 19. (Monographs of the High/Scope Educational Research Foundation, No. 8). Ypsilanti, MI: High/Scope Press.

Coleman, J., Hoffer, T., & Kilgore, S. (1982). High school achievement. New York: Basic Books, Inc.

Lazar, I., Hubbell, V. R., Murray, H., Roscha, M., & Royce, J. (1974). Summary report: The persistence of preschool effects. U.S. Department of Health, Education, and Welfare, Administration for Children, Youth, and Families, Washington, DC: U.S. Government Printing Office.

Noah, T. The fat in college budgets. New York Times. March 19, 1985.

U.S. Department of Education, Office of Special Education and Rehabilitative Services (1985). Seventh annual report to Congress on the implementation of the Education of the Handicapped Act. Washington, DC: U.S. Government Printing Office.

U.S. Department of Education, National Center for Education Statistics (1984). Quality of responses of high school students to questionnaire items. Washington, DC: U.S. Government Printing Office.

NEW AREAS FOR EDUCATIONAL DATA COLLECTION: WHAT STUDENTS ARE TAUGHT AND WHAT THEY LEARN

Elizabeth R. Reisner
Policy Studies Associates

In the next several years pressures are likely to grow for the federal government to expand its reporting of what elementary and secondary students are taught and, from the instruction to which they are exposed, how much they learn. Gathering accurate data in these two areas will present major challenges both to the National Center for Education Statistics, because of the technical problems of survey and analysis design, and to its parent the U.S. Department of Education, because of the policy debate these proposals and plans will inevitably provoke. To meet these challenges, NCES should begin now to consider approaches for gathering and reporting data on (1) the curricular content of elementary and secondary instruction and (2) students' achievement of the curricular goals that are set for them.

This paper provides an overview of topics related to federal data collection in these areas. The first section reviews the current need for such data. The second section explores how data in these areas could actually be collected. The third section surveys several problems to be addressed in designing these data collection activities.

Why We Need Better Information in These Areas

At present federal agencies provide extensive information on many aspects of American education. For example, comprehensive information is available on what have been termed the inputs of the American education system; these include the characteristics of participating students, the fiscal resources available to elementary and secondary education, the structure established to govern educational activities, and other components of the system. We also know something about the processes that affect education, including the interlocking character of governance decisions at local, state, and federal levels and the supply, assignment, and supervision of teachers and other staff. With regard to educational outcomes, we know how students score on tests that measure knowledge and critical thinking, using standard test items for students in very different schools and localities.

These measures do not, however, permit us to form generalizations about the curricular content of instruction or students' mastery of that content. For example, although we may know that high school students in a given state complete an average of three units of mathematics prior to graduating, for most states we do not know what proportion of graduating students take Algebra II, Trigonometry, or Calculus, nor do we know what mathematical skills and concepts are typically taught in each of these mathematics courses in a state. Without such information, we cannot legitimately compare the mathematics program in one state with the program of another state whose graduates typically complete fewer units of mathematics prior to graduation.

Similarly, without data describing students' mastery of the mathematics curriculum in their schools, we cannot conclude whether the completion of three units of mathematics in a state generally results in (1) mastery of a larger range of mathematical skills and concepts and (2) more thorough mastery than is achieved in a state whose students generally complete fewer units.

It is not surprising that the federal government has avoided collecting detailed data in these areas. As discussed later in this paper, it is a particularly difficult thing to do.

Nevertheless, the task is important, as illustrated in three recent reports on proposals for change in higher education (Bennett, 1984; National Institute of Education, 1984; and Association of American Colleges, 1985). The three reports look critically at the content of current undergraduate curricula and propose ways of upgrading it. Because of the sparsity of national data on undergraduate curricula, the conclusions of these reports are based on anecdotal evidence of what college curricula currently include and on small-scale surveys of course titles in selected colleges and universities. If these data were more complete, the authors of these reports would be able to make their recommendations more persuasively and indeed might find greater congruence among their analyses. Similarly, the lack of data on the content of elementary and secondary schooling precludes accurate analysis of the extent to which students are receiving instruction in skills and subject areas needed for successful transitions into these postsecondary programs.

Our lack of information on what students are taught and what they learn makes it difficult for us to analyze the fit between elementary and secondary schooling and the range of post-high school paths taken by young adults. For students going on to postsecondary schooling, college and university planners need to know what content and skills have been presented to high school students in order to design appropriate entry-level courses; in particular, postsecondary planners need to know how the content of high school programs is changing, in order to adjust entry-level undergraduate courses accordingly. For example, if recent national attention to science and mathematics instruction has actually increased students' exposure to and mastery of skills and information in these areas, entry-level undergraduate courses in science and mathematics should be upgraded to reflect these changes. For high school graduates moving directly into jobs, postsecondary vocational training, or military service, analysis of high school course content, students' mastery of it, and trends in content and mastery can indicate whether high schools are adequately preparing students for changing vocational requirements and whether changes at the high school level reflect trends in job requirements.

These data are also needed for other policy purposes. For example, they can provide a yardstick for educational agencies to compare their own programs and performance with those of comparable agencies. This information can help them identify instructional areas of relative weakness and strength. Data on instructional content and mastery can also provide a baseline for agencies to assess changes in their programs and student performance over time. This type of analysis is more useful than a static comparison of an agency's program and performance across two points in time because it permits comparison with changing national trends.

At the broadest policy level, these new data are needed to address growing demands for accountability in the expenditure of public funds. As the competition for limited tax revenues becomes more intense, legislative bodies, the media, and taxpayers increasingly require that the value of educational and other social service expenditures be concretely justified. These demands are likely to be particularly insistent in the states that have recently enacted educational reform proposals -- and new revenue raising and spending plans to implement them.

These growing requirements for programmatic and fiscal accountability provide our backdrop for consideration of how these needs can be addressed.

How to Obtain Better Data on Curriculum and Achievement

Information on curriculum is considerably easier to obtain than information on students' mastery of it. In either instance it is essential that information requests be limited to the minimum needed to yield simple data capable of meeting the needs already discussed. For purposes of designing data collection procedures, those needs may be summarized as requirements for baseline measures that permit the following:

- Comparisons across educational agencies to assess relative status
- Comparisons across time to assess change in educational agencies
- Planning by post-high school providers of education and training, including institutions of higher education, postsecondary vocational programs, and the military
- Assessment by the public to determine if expectations for educational delivery and performance are being met

The following discussion focuses first on how NCES could obtain data on the content of students' curriculum to address these requirements. Then we turn to procedures for obtaining data on the extent to which students learn the skills and information contained in the curriculum.

Instructional Content and Methods

The first step in designing procedures to collect data on curricular content is to decide exactly what information is necessary. Possible information categories include for any given subject area (e.g., mathematics, language arts):

- Courses of study offered at each grade, including which are required and which are elective
- Student enrollment in each course
- Major skills, concepts, and information taught in each course
- Local or state guidelines for skills, concepts, and information to be taught in each course
- Primary teaching techniques used in each course (e.g., laboratory or activity-centered; classroom discussion; extensive use of audio-visual materials, computers, or self-instructional materials)

- Subject-relevant qualifications (e.g., academic preservice training, inservice training, years of experience in teaching the course) of teachers teaching each course
- Role of subject supervisors or specialists in teaching (or shaping) the course

These information categories are derived from those used in a study conducted by Weiss for the National Science Foundation, entitled "1977 National Survey of Science, Mathematics, and Social Studies Education." The survey was designed to obtain data on course offerings, curriculum usage, and classroom practices in science, mathematics, and social studies. Besides taking a snapshot of instructional practice in these three subjects, NSF used the survey to determine the extent to which NSF-sponsored materials were being used in classrooms and the self-reported influence of NSF-sponsored programs of inservice training on science, mathematics, and social studies instruction. To improve the quality and comprehensiveness of the survey, an important design step was the review of the draft questionnaire by a large group of consultants with appropriate subject matter expertise and representatives of professional associations with interests in science, mathematics, and social studies education.

Another perspective on the collection of instructional data is provided by a 1981 survey conducted for the Committee on Economic Education of the American Economic Association by Yankelovich, Skelly and White, Inc. The purpose of this survey was to examine "how economics is being taught in America, i.e., who is teaching it, where it is placed in the school curriculum, what the focus is, and what teaching aids and materials economics teachers are using." As in the NSF project, a number of expert consultants and association representatives were involved in survey design and analysis. This survey collected data in categories similar to the seven listed at the beginning of this section.

The economics survey is particularly relevant to the proposal presented in this paper because it reported information on the instructional content of economics courses, using several sets of focused subcategories. For example, within the category labeled "major skills, concepts, and information taught in each course" in the list above, the economics education survey breaks out two sets of subcategories. The first is called "goals of economic education" and offered respondents the choice among six possible goals, including "to help students understand the current problems facing the country" (reported as a very important goal by 66 percent of all responding teachers) and "to teach students practical skills that they need in their everyday lives, such as balancing a checkbook, using credit cards, how to shop wisely, etc." (reported as important by 65 percent of the respondents). The survey also breaks out a second set of information categories focused on "aspects of economics"; these include 23 headings such as "supply and demand" (the most frequently taught aspect of economics) and "consumer issues/ consumerism" (an aspect taught by 66 percent of all responding economics teachers).

These two surveys demonstrate the feasibility of obtaining nationally representative data on instructional content and methods. They also suggest the desirability of conducting survey development separately for each major subject area. This strategy would permit the use of expert review panels, such as those used in these two surveys. The panels could play valuable roles in the review and adoption of subject matter categories such as those used in the economics survey. (These categories are available in other subject areas and are sometimes called "taxonomies of educational objectives.")

Students' Mastery of Instructional Content

The preceding section suggests that any testing of students' achievement of curricular goals in a particular subject area should not occur until agreement exists on a discrete set of objectives or topics for that subject area. At that point, grade-appropriate test items can be developed for each topic or objective within a subject area. Obviously, this process will need to be as careful and precise as the development of content categories, in order for the test items for a particular topic at a given grade level to meet criteria such as the following:

- Accurate measurement across a broad spectrum of difficulty levels, in order to determine the level of difficulty mastered by a student
- Assessment across the full breadth of content commonly taught in connnection with a particular objective or topic
- Assessment of differing types of achievement sought using a variety of teaching methods

Although other criteria will be necessary as well, these indicate the challenges in designing standard test items tailored to varying instructional content and methods.

Test administration will require that students be examined using the test items that correspond to the objectives or topics on which their instruction has focused. Test results under this procedure will be a more accurate measure of the effectiveness of instruction than is currently the case with the use of general tests of educational achievement.

Depending on the intended policy uses of the test results, tests can be tailored to instructional objectives at any organizational level desired (e.g., school, district, or state). For example, if a state has established improved mathematics computation as a major statewide goal, all local agencies may be required by the state to test all students in that area. Local agencies could then be permitted to administer tests of other mathematics objectives (e.g., mathematical reasoning) according to local priorities. The ability to tailor such state-level uses for the tests will be a major factor in encouraging voluntary participation in test development and administration and even cost-sharing.

Problems Likely To Be Encountered in Implementing This New Data Collection Focus

Before implementing this proposal, it will be necessary to address and resolve four sets of problems, as described below.

Problems in reaching agreement among educational interests affected by the proposal. This proposal may be seen as intruding on state and local prerogatives in the educational enterprise, because it would require educational agencies to reveal the priorities they place on different areas of the school curriculum. If widely held, this perception may require that NCES distance itself somewhat from the developmental process. One way to do this will be to rely heavily on the involvement of (1) advisors who are recognized as experts in curricular areas and (2) representatives of professional associations including both those with subject matter orientations and those who represent particular parties in the educational process. Using these groups and individuals to make qualitative judgments regarding the scope and content of necessary surveys and tests is likely to reduce public concern with any seemingly inappropriate federal involvement.

Problems in public perception of excessive burden and expense. The surveys and testing programs just described will inevitably mean new burdens on educational personnel and new public expense. The burden and expense could be reduced by use of smaller samples, although that strategy will prevent local educational agencies and smaller state agencies from obtaining information on their educational programs. The ideal arrangement would be for states and local educational agencies to find the data potentially available from these surveys and tests so desirable that they will use their own resources to pay for them just as states are currently able to contract for National Assessment for Educational Progress (NAEP) results on a statewide basis. Ultimately, however, NCES and ED will simply have to trade off the benefits of these survey and test activities against the Center's other investments.

Problems of encouraging inappropriate comparisons. Like the state protests provoked by the Secretary's "Wall Chart," data generated as a result of the proposed surveys and tests will only be able to describe a part of the educational program of a particular agency. By omitting such important variables as student characteristics, parent and community preferences, and financial resources available for education, the new data could generate misleading comparisons. Analysis of the new data will thus need to consider other, noncurriculum factors in order to avoid erroneous conclusions.

Problems of encompassing diverse educational goals. No matter how broad the participation in the developmental process, the surveys and tests cannot encompass all of the objectives, content, and methods that are used in American elementary and secondary education. Because they cannot possibly be perfectly comprehensive, educational agencies whose programs "fit" the survey structure will obtain more accurate and useful information than will agencies for which the fit is poorer. This problem suggests that the developmental process will need to continue even after full-scale implementation is under way, in order to accommodate state and local diversity and to improve the survey/testing fit across agencies.

Concluding Comments

Because of the potentially large scale of the activities required by this proposal, it is clear that NCES could not launch any part of it without extensive advance consultation and consensus-building. An early step in this process will be to identify actors in the national educational arena that have already expressed interest in these types of data collection activities and to determine what their particular objectives are. The Council of Chief State School Officers is one such major association and is in fact in a central position to advance and even implement parts of this proposal. A second early step will be to identify ways of limiting the initial implementation of the project. One way would be to confine initial survey and testing activities to curriculum and achievement in academic courses at the secondary level, since the needs and precedents there seem to be clearer than for elementary schooling. Whatever initial limits are adopted, the demand is likely to build for the types of information described here. NCES should begin now to consider approaches for addressing the demand.

References

Association of American Colleges, Projects on Redefining the Meaning and Purpose of Baccalaureate Degrees. Integrity in the college curriculum. Washington, D.C.: Author, February 1985.

Bennett, W. J. To reclaim a legacy: A report on the humanities in higher education. Washington, D.C.: National Endowment for the Humanities, November 1984.

National Institute of Education, Study Group on the Conditions of Excellence in American Higher Education. Involvement in learning: Realizing the potential of American higher education. Washington, D.C.: Author, October 1984.

Weiss, J. R. 1977 national survey of science, mathematics, and social studies education: Highlights report. Research Triangle Part, N.C.: Research Triangle Institute, March 1978.

Yankelovich, Skelly and White, Inc. National survey of economic education 1981: Grades six through twelve. New York: Author, 1981.

NEEDED RESOLVES FOR EDUCATIONAL RESEARCH

Prepared for the National Center for Education Statistics

Susan J. Rosenholtz
University of Illinois, Urbana-Champaign

Since the report of the National Commission on Excellence in Education prophesized an omen of doom for public education nearly two years ago, educational policy-makers in nearly every state have scrambled to predict and conquer educational events and practices that appear most out of control. Underlying much of the current flurry of reform activity is the assumption that the teacher's effectiveness in no small way accounts for the adequacy of student learning. From the researcher's viewpoint, the teacher workforce is indeed sorely troubled. Shortages of qualified teachers have already begun to appear in some states (NCES 1984). The intellectual caliber of new teaching recruits, at least to the extent that it is revealed by measures of verbal ability, is considerably lower than was true a decade ago (Schlechty & Vance 1981; Weaver 1981). The ability of schools to retain their most academically talented teachers, again as revealed by tests of verbal ability, is also disheartening (Lyson & Falk 1984; Pavalko 1970; Schlechty & Vance 1981). And while the success of those who remain in teaching wanes considerably after five years of experience (Katzman 1971; Levin 1975; McLaughlin & Marsh 1978; Summers & Wolfe 1977) their rates of retention in the workforce far exceed that of novices in the early stages of their teaching careers (Burlingame 1980; Charters 1970; Pederson 1970).

Because of widespread--and largely justified--alarm about the status of our nation's teaching corps, many states and localities are seeking through various means to improve their teaching forces. These efforts take many forms: written examinations for teachers, extended apprenticeship periods, financial incentives and rewards for classroom excellence, various schemes for evaluating teacher performance, and more. The plethora of interventions initiated, however, appear to rest on no solid base of valid and widely-accepted knowledge about the teaching occupation. The many attributes of effective teachers are not well understood. The sources of teacher effectiveness are even less well known. The organizational and occupational influences on teaching excellence are poorly mapped. The incentives and rewards that motivate individuals to enter the workforce, to remain teachers, and to become more effective teachers, are the subject of much conjecture by policy makers, but little available knowledge.

How can the academically talented be drawn into teaching? How can persons of ordinary ability be furnished with training, experiences, occupational conditions, and rewards that will make them more effective teachers? How can effective teachers be retained in the classroom? These are but a few of the fundamentally important questions to which policy makers need answers if the fruits of their labors are to yield a more abundant harvest in improving public education.

To resist the blandishment of well-intended but ill-informed social engineering, to provide feedback into the policy-making process so as to encourage good ideas, discourage bad ones, and permit wise mid-course corrections--these are the goals that deserve full attention and support from the research community. The gathering of systematic information about the effects of various policy changes represents nothing less than a rational basis for further decision-making, planning, and action. And with the wide variations in policy changes currently underway in states and localities, the opportunity is at hand to do precisely that.

Although sorely underutilized by policy makers, in a research sense we are already mounting a successful front against the common enemy of low school productivity. And we are equipped with sufficient conceptual, analytic, and methodological clarity, and a secure enough knowledge base, to launch further forays. In the section that follows I provide the conceptual underpinnings, but by no means an exhaustive description, of the knowledge about effective school practices that bear directly on the quality of the teacher workforce. I will look at the ways working conditions enhance teacher commitment, retention, and teaching effectiveness across a wide range of studies. Armed with this conceptual understanding, current policy decisions and their ability to affect positive changes will then be analyzed.

The Dimensions of Teacher Commitment

That the most vital resources for student learning are the contributions of effort and involvement from teachers is a proposition few would dispute. Teacher commitment and its attendant behaviors, however, are not categorical or unvarying commodities. They depend to no small extent on the incentives and opportunities offered by the school and on the organizational conditions under which teachers work. In particular, teachers are motivated both to remain within a setting and to contribute productively only so long as the inducements offered them are as great or greater than the contributions they are asked to make (Locke 1975; March & Simon 1958). In other words, the rewards of one's work must outweigh the frustrations.

Teacher rewards. There is limited information on the importance of monetary rewards in securing teachers' commitment, but the extant data provides little empirical evidence that increased pecuniary benefits bring about positive changes in teachers' performance (Mann 1985; McLaughlin & Marsh 1978), or prevent their defection from the workforce (Bredeson, Fruth & Kasten 1983; Bruno 1981a; Chapman & Hutcheson 1982; Frataccia & Hennington 1982). Teaching rewards instead flow directly from feelings of efficacy: from recognition of one's own capacities to affect student growth and development (Bishop 1977; Bredeson, Fruth & Kasten 1983; Glenn & McLean 1981; Lortie 1975; McLaughlin & Marsh 1978).

Teachers' inability to accrue psychic dividends from their work may manifest itself most dramatically in a decision to defect from the workforce (Bredeson et al. 1983; Chapman 1984; Chapman & Hutcheson 1982; Litt & Turk 1983; Rosenholtz et al. 1985). The link between dissatisfaction and actual defection, however, may be mediated by the alternatives individuals perceive to be available (Locke 1975; March & Simon 1958). A lack of alternative types of employment, for example, may cause dissatisfied teachers to stay where they are and simply withhold service. Although the particular manifestations of withheld service are not fully known, there is evidence that workers sometimes absent themselves to provide temporary relief from unsatisfactory job conditions (Johns & Nicholson 1982). Indeed, teacher absenteeism is particularly prevalent in ineffective low SES schools (Bruno 1981b; Bruno & Doscher 1981; Spuck 1974), where large discrepencies sometimes exist between the inducements of teachers' work, and the contributions they are expected to make.

Teacher certainty. In addition to psychic inducements, productive involvement in work also requires challenge (Locke 1975). Challenge stimulates involvement by requiring that individuals exercise judgement and choice; in doing so, they become the main causal agents in performance. Further, coping with challenge requires the expenditure of effort. If this expenditure produces some improvement in performance, commitment is enhanced. Individuals move initially toward confronting challenge, however, only when there is a reasonable chance of success--some assurance that their efforts will produce desired outcomes (Campbell & Pritchard 1975). In the case of teachers, commitment to meet classroom challenges pivots fundamentally upon their certainty about professional practices--a belief in their ability to help students grow and develop (Azumi & Madhere 1983; Glidewell et al. 1983; McLaughlin & Marsh 1978; Rosenholtz et al. 1985). When certainty pertains, it defines and organizes teacher action to facilitate student learning (Armor et al. 1976; Ashton et al. 1983; Brookover et al. 1979; McLaughlin & Marsh 1978). The other side of the same coin is that challenges perceived as too great or costly may cause individuals to experience a sense of failure and frustration, leading often to inaction (Locke 1975). In other words, teachers who are uncertain about their capacity to affect student learning tend not to act in ways that will bring learning about. One need only consult the plethora of research on differential teacher expectations to see how powerful this self-fulfulling prophecy can be (for a review, see Hawley & Rosenholtz 1984). Because the products of uncertainty--e.g. low student learning and teachers' sense of failure--ultimately diminish teaching rewards, it is not surprising that teachers who lack confidence in their professional skills tend to show higher rates of absenteeism and defection from the workforce (Chapman 1984; Chapman & Hutcheson 1982; Litt & Turk 1983).

Skill development. To secure individuals' commitment, the work setting must not only provide challenges, it must also provide opportunities to deal successfully with them (Locke 1975). It follows therefore, that opportunities for skill acquisition and development that enhance teachers' capabilities are heavily implicated in their commitment. While there is a dearth of research on this assertion, the significance of skill development for disaffection seems logical enough: limited opportunities for professional growth impair teachers' certainty about instructional practice, their effectiveness, their acquisition of intrinsic rewards, and ultimately their commitment to the school and profession (see, for example, Huberman & Miles 1984; Rosenholtz et al. 1985). Not unexpectedly, the absence of opportunities for professional growth is frequently cited by teachers as a reason for disaffection and attrition (Bredeson et al. 1983; Mann 1985; Rosenholtz et al. 1985).

The three intervening variables affecting commitment--teacher certainty, skill acquisition, and rewards--rely heavily on the actions of others within the school--colleagues and principal--and are thus strongly influenced by specific organizational policies and practices. I turn next to a description of these additional factors.

Organizational Determinants of Teacher Commitment

Teacher evaluation. Skill acquisition, certainty, and rewards depend to no small extent on feedback about one's performance--for teachers, on evidence of student growth and learning. Yet teachers frequently indicate difficulty in knowing precisely how well they are doing (Ashton et al. 1983; Glidewell et al. 1983; Lortie 1975), leaving many uncertain, unrewarded, and without the specific information needed to redirect their energies toward improvement. Ambiguity about the nature of one's performance springs at least in part from an absence of both clear goals around which to mobilize teaching efforts, and clear criteria by which teacher performance is monitored and evaluated.

While many school administrators muster little effort to resolve this ambiguity for teachers, those in the most effective schools develop clear goals and ubiquitously monitor classroom efforts toward their pursuit (Glenn & McLean 1981; Hort, Steigelbauer & Hall 1984; Natriello 1984; Natriello & Dornbusch 1980; Sizemore et al. 1983; Venezky & Winfield 1979). With clear, useful, and frequent evaluation, teachers can work directly to improve performance; as performance improves, there is greater certainty about instructional practice, and with it renewed teacher effort and larger psychic dividends (Rosenholtz et al. 1985). It is not surprising, therefore, that teachers report greater satisfaction and commitment where principals provide frequent and clear evaluation (Azumi & Madhere 1983; Chapman & Lowther 1982; Natriello 1984; Natriello & Dornbusch 1980; Rosenholtz et al. 1985).

Buffering. Greater commitment is also secured by working conditions that facilitate individuals' attainment of work goals (Locke 1975), for teachers, on conditions that optimize the possibility of student learning. Intrusive managerial tasks that pull teachers away from instruction are frequently culpable in the absence of their skill acquisition, certainty, rewards, and commitment (Bredeson et al. 1983; Lortie 1975; Raschke, Dedrick, Strathe, & Hawkes 1985; Rosenholtz et al. 1985).

Efficacious principals (or their administrative cadre), themselves certain of the relationship between teacher effort and student learning, work to "buffer" teachers from unnecessary intrusions that distract them from the substance of their work. Buffering activities include attending to the material requirements and organization of instructional programs, providing clerical assistance or outside resources for routine, non-teaching tasks, and protecting classroom learning time from interruptions such as loud speaker announcements and other low priority matters (Armor et al. 1976; Glenn & McLean 1981; Hort et al. 1984; Rutter et al. 1979; Venezky & Winfield 1979).

Managing student behavior. Effective administrators also distinguish themselves from their ineffective counterparts by setting and enforcing clear expectations for student behavior (Brookover et al. 1979; Glenn & McLean 1981; Rutter et al. 1979). A climate of disorder does more than frustrate teachers; when teachers attend constantly to mediating classroom disputes, they do so at the expense of their students' learning time, their own instructional improvement, their confidence about teaching skills, and any psychic rewards that follow (Raschke et al. 1985; Rosenholtz et al. 1985). This explains why teachers often cite student misbehavior as a cause for dissaffection and attrition from the workforce (Bredeson et al. 1983; Raschke et al. 1985; Rosenholtz et al. 1985). And since learning to manage student behavior is the first important task of the teaching neophyte--and one that is used as an initial measure of their potential (Hoy 1969; McArthur 1979; Warren 1975)--it is not surprising that attrition is highest in these early years of teaching.

Teacher isolation. Most schools are characterized by isolated working conditions, where colleagues seldom see each other teach (Bishop 1977; Cohen 1981; Lortie 1975). Under these conditions norms of autonomy develop, where the responsibility for classroom success resides solely with individual teachers (Bishop 1977; Glidewell et al. 1983; Lortie 1975). Requests for and offers of assistance among faculty are believed to carry status information about relative teaching competence: teachers tend not to request assistance for fear of appearing incompetent; teachers tend not to offer assistance for fear of implying incompetence (Glidewell et al. 1983; Lortie 1975). We know little about how such faculty norms develop, but it is reasonable to suppose that the primarily social conversations that characterize teachers' interactions in isolated settings (Bishop 1977; Glidewell et al. 1983; Little 1982) occur in an effort to avoid these status implications.

Professional isolation has a profound effect on teachers' skill acquisition, certainty, and intrinsic rewards. For one thing, their capacity for growth is limited in isolated settings by their own ability to diagnose problems, develop solutions, and evaluate their effectiveness (Lortie 1975). With little access to role models among their peers (Gehrke & Kay 1984;-Lortie 1975), they realize little benefit from their more experienced and expert colleagues. Similarly, the intrinsic rewards to be derived from colleagues' positive evaluations of one's skills and ideas are foregone in isolated settings (Rosenholtz et al. 1985).

Faculty collaboration. Not all schools are isolated workplaces; in the more collaborative settings of effective schools, teachers come to believe that teaching is a collective rather than an individual enterprise. Professional dialogue among colleagues in these schools is frequent, and analysis, evaluation and experimentation with colleagues set the conditions under which teachers improve instructionally (Armor et al. 1976; Little 1982; Mann 1985; Rutter et al. 1979; Venezky & Winfield 1979). In collaborative settings, teachers interact more about professional than social matters, and interact with a greater number of colleagues than is true in more isolated settings (Bishop 1977; Bridges & Hallinan 1978; Glidewell et al. 1983; Little 1982). Ideas that are the product of collaborative exchange appear to give rise to greater experimentation in classrooms, and greater teacher learning and certainty, as better solutions to teaching problems are found (Rosenholtz et al. 1985). And it is precisely these conditions that most clearly explain why teacher absenteeism and defection are substantially lower in collaborative than in isolated settings (Bridges & Hallinan 1978; Litt & Turk 1983; Rosenholtz et al., 1985; Sizemore et al. 1983; Venezky & Winfield 1979).

We know therefore, that informal learning experiences can and do influence teaching knowledge. Strikingly absent for the literature, however, are studies that look at the combination of formal and informal mechanisms that shape teachers' beliefs about the definition of "what good teaching is," that accentuate or enhance the acquisition of skills, that define the standards by which teachers measure their success in teaching, and that therefore signal the need to develop new teaching skills or perfect old ones.

Participation in Decision-Making. One informal mechanism that may account for collaborative exchange among faculty in efficacious schools is teachers' participation in decision-making about matters related to teaching, e.g., selecting instructional materials and methods (Armor et al. 1976; Glenn & McLean 1981; Rosenholtz et al. 1985; Rutter et al. 1979). The contribution of decision-making to teachers' skill acquisition and certainty lies in the deliberative evaluation, discussion, suggestion, and modification of instruction required to enhance the quality of classroom learning. These activities themselves may lead to increased teacher clarity about instructional purpose, and ultimately, to greater effectiveness, as decisions become conscious and well-reasoned choices rather than arbitrary or automatic reactions (e.g. Cruickshank 1985; Mann 1985).

In addition to its effects on skill acquisition, certainty, and intrinsic rewards, participation in decision-making may directly augment commitment through an increased sense of school ownership as teachers identify their own important contributions to a valued collective enterprise. The point here, of course, is that teachers who do not subscribe to the faculty's purpose are not likely to contribute their full efforts. This explains most clearly why the absence of teachers' involvement in decision-making is positively related to their absenteeism and defection (Azumi & Madhere 1983; Chapman & Hutcheson 1982; Rosenholtz et al. 1985).

Organizational rigidity and flexibility. Another informal mechanism that is implied by norms of collaboration and decision-making is the organizational flexibility necessary to alter instructional programs to meet specific classroom needs. However, uncertainty about the ability of teachers to help students learn may sometimes lead principals to apply excessive pressure for conformity to rules and regulations which may themselves be overly specific. Insistence on ritualistic adherence to school procedures not only leads directly to profound teacher disaffection (Hoy, Tarter & Forsyth 1978), it also produces greater feelings of powerlessness and uncertainty (Cox & Woods 1980) and diminished focus on learning goals (Willower & Jones 1963). In the end then, organizational rigidity reduces teachers' ownership of instructional programs and pays few dividends in teachers' skill development, psychic rewards, or commitment (Raschke et al. 1985; Rosenholtz et al. 1985).

Education Reform

In the preceeding discussion I have emphasized the importance of workplace conditions in helping teachers develop, perfect and add to their fund of teaching skills throughout their professional lives. This conceptual understanding provides a framework to gauge the educational efficacy of many reform proposals that have gained considerable currency among policy-makers. The central question in all discussions of school reform is, of course, the extent to which changes will improve the quality of instructional services that schools deliver. How can schools be restructed in ways that permit teachers to use their talents most productively in helping themselves and others to improve instruction? How will new standards for student learning affect what teachers actually emphasize in their classrooms? How will the performance recognition and status elevation of more talented teachers as proposed in various career ladders plans affect both their motivation and commitment as well as that of others? Under what conditions are these innovations most likely to succeed?

While any number of reform efforts might be juxtaposed against our cumulative knoweledge, I examine the effects of two specific proposals--standards for student learning and career ladders--to illustrate the rather straightforward proposition that through careful research and analysis, sound information about intervention strategies supplied to policy makers will enable them to make sense out of the many events that are occuring so that they can be helped to foster and support further school improvement.

While any school intervention can be developed, funded, supported and delivered to schools to help them improve, the ultimate measure of their success depends to no small extent on how the intervention is executed. In this the teacher is crucial. How teachers perceive and experience policy changes will affect their commitment to them and the extent to which the policy change will have a salutary effect on student learning. To explore teachers' perceptions, I turn to data from our ongoing study of the organizational conditions of teaching (Rosenholtz, Bassler & Hoover-Dempsey 1985) conducted in a Southeastern state where a career ladder plan (CLP) and minimum competency testing (MCT) are currently under implementation. Data from extensive interviews we have conducted with 73 randomly selected elementary teachers statewide will illuminate many of the issues I raise. I will also draw upon contemporaneous work by others who seek to chronicle and understand the effects of reform efforts underway elsewhere in the nation.

STANDARDS FOR STUDENT LEARNING

Some states have begun to recognize past inadequacies in the way student performance is monitored, and attempts are now underway to institute appropriate changes. One such strategy sets skill-specific standards and learning sequences that all teachers must cover in their classroom curriculum. Conformity to new standards is monitored through the periodic and frequent testing of students.

The Importance of Standards

Of fundamental importance to any policy study of educational reform is the definition of student learning itself. In isolated settings, teachers' and administrators' perceptions of student learning are highly individualistic, since they are based on those classroom activities and student behaviors that each considers important. They may include students' problem solving skills, peace and quiet in the corridors and classrooms; the development of youngsters' self-concept; children's basic skill acquisition; the inculcation of racial tolerance and friendly interpersonal behavior, and so on.

Yet the literature on effective schools apprises us of the importance of shared organizational goals (Rosenholtz 1985a). Where there are particular goals for students' basic skill acquisition, agreement among teachers and administrators as to their importance, and collaboration about the means by which to implement them, there is an organizational basis for directing teacher behavior, for motivating teacher behavior, and for evaluating teacher behavior. Goals, then, can be useful in mobilizing the efforts of school personnel by providing specific targets and directions for change.

In the southeastern state we have been studying, the department of education established MCT for elementary grade students, by identifying 1,300 skills in reading and math, of which 680 must be learned. The interviews we conducted with teachers two years after implementation revealed uniformly high conformity to the guidelines as well as accompanying changes in teachers' instructional emphases.

Twenty-two percent of the teachers interviewed found the guidelines helpful in detecting students' learning difficulty, and, of far greater significance, in changing the way learning is perceived by both students and teachers. As one explained, "I find skills that children don't have and so I have to teach them. Now teachers care that students learn. Before they could just teach and if the students learned, okay; if not, they could just go on to the next thing. The kids also know that 'I have to know this' and 'my teacher cares that I get it.' So it helps." Others concurred with the meliorative effects of MCT on poorer teachers, e.g. "I think the standards are more effective for teachers who need guidelines." In fact some teachers hailed the change as a way to orient their own classroom instruction, thereby ensuring that the most important skills receive adequate time and attention: "I think it helps a teacher measure her own teaching. Lots of times I will compare my tests with their [the state's] guidelines. Now I'm more aware of some specific skills and how well I have taught them as well as how well they're covered in the curriculum." Darling-Hammond and Wise (1985), in their study of evaluation practices in two middle-Atlantic states, also found that a minority of teachers regarded MCP as a good management tool for helping less competent teachers to do their jobs, and for ensuring that a specific body of knowledge was covered in the classroom curriculum.

A critical factor in the study of state level intervention, then, at least at the elementary school level, is the extent to which standards alter the goals that teachers' set for themselves, what they come to emphasize in their classroom curricula, and how their teaching effectiveness is gauged. That is, where standards for student evaluation are clearly specified, and where teachers may also be judged by their students' abilities to reach these standards, instructional content may become driven by newly implemented standards and their measurement. Anecdotal evidence suggests that this may be already occuring in Texas, Detroit, South Carolina; and Maryland (Popham, Cruse, Rankin, Sandifer, & Williams 1985).

Unintended Consequences of MCT

Organizational rigidity. MCT, however, was not uniformly welcomed by all teachers, and there were some unintended and negative consequences in its implementation. One concern, voiced as well by teachers in the Darling-Hammond and Wise (1985) sample, was that rigid standards impaired teachers' discretion to match appropriate learning objectives to particular student needs. Over one third of our sample expressed this objection, e.g., "All kids are to be exposed to all skills. I don't think that does any good. If you progress too fast, the kids still lose out. The kids have to learn the basics first, or it won't do any good to expose them to other skills"; "Sometimes I give tests at different times than I'm supposed to. I'll teach the skills first, and then give the test, regardless of when they say I'm supposed to give the test. It doesn't make much sense to test kids on stuff they haven't studied"; "It used to be that we would take up where the kids were and you would go as far as you could with them. Now they say we are not to do that anymore"; "The business of all second grade teachers teaching the same across the state, being on the same unit, the same page, is the most absurd thing I have ever heard in my life. We are all supposed to be teaching the same thing at the same time so if a child transfers from one school or system across the state, they won't be behind"; "There are too many skills to get through when some children can't even regroup. Because children are different you have to make allowances for them"; "With the Basic Skills Program everyday I feel more frantic. I feel like a cattle-driver with a whip: like we have to get through the pass before nightfall."

In the implementation of MCT, then, teachers confront a dilemma between the coverage of required basic skills on the one hand, and rudimentary mastery of them, on the other. Thus a critical question in the study of MCT is the extent to which it allows for local variations within students' skill levels, and local deviations from statewide norms. To inhibit the appropriate pacing of instruction to accomodate different learning needs or to prevent the adaptation of curricular content to improve its fit to those needs, is to unwittingly program students and teachers for greater academic failure.

Another quarter of the teachers in our sample complained that curricular areas other than reading and math were being short-changed by state standards. Even kindergarten teachers lamented these effects: "I am not able to do things that are good for kindergarteners. I feel like I have to hide in my room to let children have show-and-tell"; "I wanted to do more creative dramatics and storytelling [this year]. I wanted to expand my study of marijuana that I instituted last year. Maybe draw in some teenagers to talk with the class. But there is not a lot of time to do anything like that. I did the drug unit during health but I had to steal time for the dramatics." Darling-Hammond and Wise (1985) also found that the need to ensure that their students pass competency tests caused teachers' to de-emphasize other important aspects of the curriculum.

Lest readers doubt the value of alternate learning opportunities, they need only consider the effects of MCT on language arts curricula alone. Writing instruction and practice in some classrooms have been replaced by rote exercises in sentence diagramming (Suhor 1985), an ineffective instructional strategy in helping students better their writing skills (Sherwin 1969), but nonetheless content most likely to appear on competency tests (Suhor 1985). It is indeed an unfortunate side effect if students' opportunities to master a broad base of knowledge are undermined because teachers divert their instructional emphases either to material that is to be tested, or, worse still, to the teaching of test-taking skills themselves. This latter charge was alleged by Los Angeles school board member and teacher Jackie Goldberg: "Teachers who used to spend time reading to children so they will love literature will now have them bubbling in dots [filling in computer cards] of how basesballs are stitched because this is one of the questions on the standardized tests." (Chicago Tribune, June 2, 1985, p. 13).

A third of the teachers we interviewed balked at the standardization demanded by the MCT which in their view destroyed teacher creativity and spontaneous teaching, e.g., "Teaching has become more mechanical; [the MCT] took away creativity and the teacher's individuality"; "The system has become very dictatorial. I mean I ask myself all the time, 'I wonder where all the creative teachers will be in ten years'"; "Twenty years ago they would say to me, 'As long as you get the skills taught, do it your own way.' We are as individual as the children are. And what works for my next door neighbor doesn't work for me. It's amounting to having all teachers teach the same way."

The absence of buffering. By far and away, the most onerous aspect of MCT to three-quarters of the teachers in our sample was both the overwhelming burden of additional paperwork, and the classroom time required to test. Over both these points there was nearly unanimous accord--e.g., "You have to assess and reassess each child, you have to pre- and post-test each child. I want to keep good records; I always have kept good records, but things are just getting out of hand." A kindergarten teacher complained, "Each student has to take 20 to 30 tests. It seems to be that basic skills is more testing than teaching. That is all I do. Each test takes about 15 minutes per child. I have to give the tests to each child one at a time. There just doesn't seem to be any benefit in the program." Indeed, teachers we interviewed reported that valuable instructional time--the teacher's most prized resource--was considerably diminished by paperwork and testing demands--e.g., "I am actually teaching less"; "There is too much testing rather than teaching"; "I'm just not sure the kids are learning basic skills"; "It takes away from actual time spent working with children"; "There's not much teacher time to be human towards the students". Teachers chronicled for us precisely how they accomodated new paperwork demands by reducing their instructional time with students: "I really feel bad because I'll let the kids have five extra minutes of play or give them independent seatwork so I can get some of my work done. I feel bad about taking time away from my students, but I have to."

Where daily planning had once occupied teachers' after school hours, record keeping now takes its place. Where teachers once interacted before school hours, myriad state forms now compete successfully for their attention. Again typical of their sentiments are these teachers' comments, "With the paperwork, many teachers stay until seven or eight p.m. But if you have a family, you can't do that. What I want to know is, when are we supposed to teach?". "I feel like I'm robbing Peter to pay Paul. The time has to come out of somewhere, doesn't it? I can't not sleep each night because I have to do paperwork for the state. So I have to take it out of my teaching time." Rather than giving students greater opportunity to learn basic skills, and testing them to ensure mastery, MCT may instead rob them of access to their most critical learning resource--teachers' instructional time. Indeed, Darling-Hammond and Wise (1985), report identical concerns issued by teachers they interviewed. Most revealingly, teachers in our sample that found benefit in MCT reported being adequately buffered from the additional paperwork and testing demands by outside clerical assistance from either paid aides (provided by the district) or parent voluteers (coordinated by the building principal). For these teachers, then, MCT provided the intended and welcomed feedback about student progress that could serve to redirect their own teaching strategies.

Lowered teacher commitment. Implementation of MCT in the majority of schools may ironically also cause teachers to feel professionally violated: "The amount of paperwork takes all the fun out of teaching. What really bothers me is that the teachers' judgement is not considered important any longer. We used to be able to decide things...Now we teachers are frustrated. Every new program puts new burdens on us. Every new program means less time to do work with actual teaching and being with the kids. We dislike that a lot." Teachers perceived that new policies have been enacted with little understanding of teaching realities, e.g., "I can't stand it when they make decisions about what I'm supposed to be doing when they don't understand what teaching is all about." And this: "For someone else to tell me what they think is needed when I can see some other things that are needed myself is infuriating."

Twenty percent of our sample either openly contemplated leaving the profession, or reported others were doing so, because of the increased testing and paperwork demands and their inability to derive psychic benefits because of it: "I have enjoyed teaching but I am planning to retire early because I have been frustrated in my ability to do what I know is best in my own classroom. I think that with the amount of paperwork that we have, the recording and testing and everything, that I'll leave that to someone younger." "I am really afraid that all the good teachers are going out of teaching. Sometimes I ask myself, 'Will we only have desk-sitters in the future?'"

Over sixty percent of our sample of teachers complained of lower morale on their faculties brought about by MCT: "Everyone feels bad about being a teacher these days. Everytime you turn around it seems there's always someone telling you that you're doing a lousy job. I've stopped reading the newspaper at all because if they say something negative about teachers it doesn't make me want to go to work that day... All we [the faculty] seem to do these days is complain. We all feel it. I don't think the public realizes how hard it is to do a good job--to work through teaching problems--when everyone's saying you can't."

"I think the morale of teachers is very low now. The teaching load as far as bookwork, paperwork, is just weighing them down so heavily that they resent spending their time with paperwork and not actually teaching. If we had aides to helps us put it on a computer, then we could spend more time teaching. We did have one aid for the Basic Skills Program, but she was spread so thin that she was just not that helpful to any one teacher. Teachers just realize that there are not enough hours in the day, so many of us will bring stacks of work home, and I work almost every night until 8:00 or 9:00 o'clock, sometimes midnight, at that gets old after a while. You have to enjoy what you're doing, and I do enjoy being with children, if I just had more time to teach them instead of filling out reports."

Some teachers described the personal, human costs that the decline in morale brought on: "You just don't have the enthusiam you once had. There are times when maybe I have not been as patient as I should have been. You seem to get wound tighter and tighter. There was a particular time when I really lost my temper. I was so frustrated because I had worked so very hard with a child--I would stay after school and work with him and sometimes I would even take him home. And then it was just sort of a let down...You feel like you are beating your head against the wall. I'm coming to the point that I'm enjoying teaching less. It's becoming more of a job instead of something you want to do. If I thought I could get out of teaching and into something else, I would. I've heard that from a lot of teachers. Sometimes you can put up with certain things, but when things begin to overwhelm you, everything seems to just drag you down. Most of the teachers I talk to just hate to go back to school--that's not the right attitude. It's such a frustrating thing to see really good teachers just turned off. A lot of people, rather than go through the hassle, just let things slide. I think this is what's happening. They don't resist, they just give in to it...I don't think it's only me, I think it's a lot of teachers. And I discourage every child that that comes back to talk to me from going into teaching. I tell them there's no future in it. That's the way I feel right now." Said another, "In the past I've always enjoyed teaching. I felt like I helped in some way. Now there is so much other than teaching I am required to do. I guess I am just burned out. I am not looking forward to the Fall."

That teacher commitment may be reduced by policy changes is another noteworthy labyrinthine of education reform that begs research disentanglement. If policy changes pose too great a burden, teachers may disinvest from their work, "just let things slide" and receive social support from colleagues for the divestiture. The possibility that increased demands which teachers perceive as barriers to their classroom effectiveness may cause good teachers to defect must be entertained and examined. In sum, researchers who chart policy changes need be mindful of this fundamental paradox: The administration of MCT may place new demands that create additional problems--lower teacher commitment--that worsen the very instructional services the reform effort intended to improve. Problems that arise from the implementation of new policy are, of course, not intractable. But without research activity that assesses the effects of policy change on the teacher workforce, and without proper procedures that feedback essential information and recommendations to policy makers, there will be no corrective action undertaken.

CAREER LADDERS

Career ladders--a proposal to reward and encourage teaching excellence--offers teachers increased salaries and status in return for taking on additional school-system responsibilities. Well-conceived CLPs (e.g. Charlotte-Mecklenburg) intend to bring about a salutory effect on schools through functional assignments in which talented teachers help their colleagues improve. Functional activities include the clinical supervision of probationary and experienced teachers, and the conducting of school inservice programs. The benefits of functional assignments clearly lie in their potential to mold schools into highly collaborative environments.

The Potential of Functional Assignments

Mediated entry into teaching may help beginning teachers, a group that defects most frequently from the workforce. Where novices receive no guidance from experienced, successful teachers, they undergo severe "reality shock", as idealism yields to the understanding that before one can teach, it is necessary to manage students' sometimes unruly behavior. In isolated settings, reality shock prompts rather negative work orientations. The view that each student has different needs gives way--usually within the first year--to a custodial view where the maintenance of order is stressed, students are distrusted, and a punitive attitude toward control prevails (see Ashton et al. 1983; Bishop 1977).

New teachers in collaborative settings, however, appear to maintain the view that tending to the individual needs of students is essential (Ashton et al. 1983; Bishop 1977). The emphasis on skill development in managing student behavior helps beginners avoid a custodial attitude, which in turn lessens their reality shock. Thus, with mentoring by highly skilled teachers, beginners' disaffection and subsequent defection from workforce may decline substantially.

Supporting the work of novices benefits experienced teachers as well, because the challenges and problems provide greater opportunity for public recognition and skill utilization. Indeed, in collegial settings, veteran teachers are more likely to perceive themselves as influential and skilled than teachers in isolated settings (Ashton et al., 1983; Chapman & Lowther 1982; Cohen 1973). Providing teachers with the opportunity to assume responsibilities, initiative, and authority commensurate with their talents and abilities, and recognizing them for a job well done, may increase their psychic rewards and their likelihood of remaining in the workforce (Chapman & Lowther 1982; Frataccia & Hennington 1982; Rosenholtz et al. 1985).

Of even greater importance to experienced teachers, however, is the degree of professional support they receive from colleagues. Beginners develop initial skills by trial-and-error learning and begin to deplete their personal fund of ideas after about the fifth year of teaching (McLaughlin & Marsh 1978; Summers & Wolfe 1977). It is at precisely this point that the organizational conditions of teaching become most crucial. Indeed, comparing the effects of school organization on relative newcomers who had taught from between one to five years with veterans who had taught from between ten to fifteen years, Rosenholtz & Greer (1985) found that organizational conditions explained 60% of how much beginners report learning, but 72% of how much veterans report learning. For experienced teachers particularly, a repository of ideas, techniques, and models, like a centripetal force, pulls them toward the same mission of professional improvement so essential to their continued commitment to the profession.

What is the potential for career ladders to develop collaborative arrangements in schools? Its success, of course, depends on how carefully the CLP is designed and implemented. Hart (1985) instructfully details one district's attempt to institute a CLP from 27 interviews she conducted with the district's principals, teachers, and superintendent.

Within this Utah district, the superintendent, assisted by a task force of administrators and teachers from each of the schools, developed a plan aimed at improving instruction by marshalling the resources of experienced and talented teachers for school-wide curriculum and instruction improvement efforts. Explicit in the plan was a commitment to the individual school as the most promising organizational level for improvement and change. Ideas were carried back and forth to faculties through task force representatives. By negotiating rather than mandating the plan, teachers developed a sense of ownership. Indeed, at the time of its implementation, 80 percent of the district's teachers voted in favor of it.

The career ladder consisted of four steps. The two highest levels—teacher specialist and teacher leader—carried with them $900, plus pay for additional contract days to work on instructional improvement projects, clinical supervision, mentoring, and assisting probationary teachers with professional development. Several benefits accrued to schools during its first year:

1. During the extended contract days, planned opportunities for teacher

collaboration were organized which resulted in increased faculty interaction and group cohesiveness.

2. Teacher leaders provided inservice based on topics identified by faculties. Teacher specialist roles were defined by each school and their number allocated by school size with an eye toward serving specific faculty needs (e.g. the number of probationary teachers needing assistance and supervision, specific program needs, faculty expertise, etc.).
3. Probationary and other teachers began to request technical assistance _on their own initiative_ from teacher leaders who had been selected because they were esteemed colleagues. Teacher leaders benefitted a great deal from these interactions as well.
4. Teachers at _all_ levels received reinforcement for the quality of their work. Teachers gained more knowledge of their colleagues' skills and talents.
5. Because teacher leaders were empowered and legitimized by their expertise, they shared (albeit sometimes in intimidating ways) decision-making responsibilities with building principals. As one teacher leader explained it, "There are nine people in this school who, in addition to the principal, think about the whole school and how to improve it" (Hart 1985, p.9). Principals and faculties confronted and communicated with each other on professional issues; faculty meetings evolved into substantive decision-making arenas.

In addition to the many structural features that accounted for the district's success is the unwavering leadership of the superintendent. It should be noted that at the onset of the plan, the superintendent of four years had already put in place the ingredients of an effective district (Murphy & Hallinger in press). For instance, the central office supported and encouraged teacher growth by bringing in resource people to work with teachers and by implementing a clinical supervision model for principals. In improving school quality, others have noted that the commitment, involvement, and active support of the central office is pivotal (Clark, Lotto & Astuto 1984; Huberman & Miles (1984), Hallinger & Murphy 1982). Certainly a superintendent who initiates experimentation and change in his district sets the tone, invitation, and expectation that others will do likewise in their schools.

Unintended Consequences of CLPs

We also find unintended and negative consequences in the forging of CLPs, and among their many variations there is grist for the policy researcher which, if combined with a mechanism for feedback, results in guidance for the reform itself. Some of the problems states and localities confront in their efforts to implement CLPs are identified below (see also Rosenholtz 1985b).

Evaluation standards. States and districts can and are identifying evaluation criteria that, because they are based on the teaching effectiveness literature, will probably differentiate effective from ineffective teachers. Careful validation studies still must butress local measures. However, the challenge to devise means that distinguish competent from great teachers has not, it appears, been successfully confronted. Indeed, in the districts studied by Hart (1985) and by Natriello & Cohn (1984), the evaluation procedures plagued both localities for their lack of clarity and definition. If exceptional teaching remains more a reputational than an observable phenomenon, the implications for changing good into great teachers are few. Moreover, as we shall see below, the success of CLPs is jeopardized if teachers do not accept new testing procedures as a legitimate gauge of their classroom effectiveness.

The Southeastern teachers we interviewed were confronting the state's evaluation and selection procedures for career ladder advancement for the first time. Nearly two-thirds of them challenged the fairness and legitimacy of the evaluation system. For example, over half the teachers charged that the classroom observation procedures and the materials submitted by the career ladder applicant measured teacher cunning and endurance more than their effectiveness, e.g., "You can do anything for a few days if you know an evaluator is coming in"; "The evaluators only make three observations, two of which are arranged with the teacher. You can fool anybody for a couple of days"; "I was observed with the criteria they use in the career ladder. I thought that was kind of farce because you know when they're coming. They don't see the true teacher"; "Anyone can put on a good show when they are being evaluated"; "You can't judge a teacher by three thirty minute visits"; "The evaluators who came in were nitpicking. They were looking for picky things. They [the teachers who applied] would tell me things they got marked down on, they got real discouraged, real uptight, a lot of them dropped out, and the ones who went on were very depressed because they didn't make it. I just had the feeling that there weren't any funds and they had to make it hard, hard."

The applicant for the two highest levels of this CLP is require to submit an astonishing array of background materials such as sample lesson plans, behavioral objectives, and teacher-made materials, which apparently was weighted more heavily than actual classroom observations. According to teacher reports, these may be fabricated without the dimmest glimmer of relevance to one's actual classroom performance. Typical of their comments: "You have to write lesson plans, unit plans, and document everything with letters. People can really make this stuff up if they want to. I know people who are doing that. The main thing is that it doesn't show whether you're a good teacher or not"; "A person who is a good test-taker and does well assembling material could be a rotten teacher. They could fool anybody"; "Just because you can write down beautiful words in a portfolio does not make you a Master Teacher. I've heard talk that some teachers are getting others to write their portfolio. I don't know. I can't prove that. But you can always get somebody to do something for you for a price."

Worse still is the pervasive complaint that the construction of the portfolio robbed students as well as family members of applicants' time and attention: "The hours needed to develop a good portfolio do not reflect a good teacher, so I decided to drop out. I also think that all those hours take away from the children. The teachers don't go in fresh. Only one teacher in our school stayed in"; "The teacher across the hall was there 'til 7:30 at night. I mean for months. You have all you can do to teach school without having this extra burden on you"; "A friend of mine who is a very good teacher applied. She felt like she had neglected her child because this had taken so much time; her weekends, hours upon hours of things that weren't really applicable to great teaching"; "A good friend of mine applied for career level three. She gave her whole year to three. Not only did it take away time from her classroom; it took away from her six-year-old duaghter, who finally one night begged her mother to just take the time to talk with her, because she spent every waking minute working on that stupid portfolio, and all that junk that they wanted in there. She had to choose one or the other".

Repeatedly teachers stressed that either they or their best-performing colleagues chose to devote their year to students rather than to developing a portfolio. "There are some extremely good teachers who are not going to apply simply because of the time factor that the others have gotten involved in"; "My principal really encouraged me to apply. He said that I could do it, and that I was already doing this and already doing that. And I did sign up for it. And then they started making changes and adding this and adding that, and you had to go take this test over, or go take this test in addition to that test...That was taking away from my time, and I just felt that I really needed the time to work on things to help my children and things for my classroom. So I dropped out"; "The time I spend on my job will be spent on preparing cleasses not on a portfolio, or running down the hall asking people to sign papers that I have had a student teacher or a field trip. I use my time and energy on my job"; "So many of the teachers who are excellent didn't apply because they just don't have the time. If you did lesson plans the way they wanted, you wouldn't have a home life."

Distributive justice. That career promotions may be based on faulty evaluation practices may stir teachers' sense of injustice. If the procedures by which the distribution of rewards are preceived as unjust or unfair--i.e., if the contributions of rewarded teachers are perceived as no greater than those of the unrewarded--problems of distributive justice arise. Unrewarded individuals react to injustice by attempting to restore equity in the setting. Typically they may alter the level of their own contributions downward in the direction of lower productivity, or they may leave the situation altogether (Cook & Hegtvdt 1983).

Nowhere is the theory of distributive justice better illustrated than in Natriello and Cohn's (1985) case study of one school district's efforts to implement merit pay. Here the Board of Education eliminated all across-the-board pay increments, using only competitive merit increments to raise teachers' salaries, a decision that brought scrutiny to the accuracy of evaluations. According to some teachers, the evaluation system forced principals to focus on relatively trivial aspects of teaching in order to make performance distinctions. Many teachers could not understand what they could possibly do to improve. And because they received only average increment raises instead of maximum raises, teachers who once felt superior now felt as though they were not performing adequately. The end result for some, as the theory of distributive justice predicts, was a reduction in teacher commitment.

In our study, teachers' persistent challenge to the soundness of evaluation practices caused many to forebode trouble when rewarded teachers begin to make substantially higher salary than others. Typical of their comments were the following: "I really don't feel like the teachers who applied [to the CLP] are doing a better job. There's going to be a lot of conflict"; "Neither of the two teachers who've applied for the top career levels are the best teachers in this school. If they make it, the rest of us will resent it terribly."

We interviewed some teachers just after career ladder selections had been publically announced. Each expressed grave reservation and surprise about at least some of those chosen to advanced: "I know someone who just got to career ladder position two and she's one of the poorest teachers I know. I like her as a person, but she is a lousy teacher"; "I think the Carrer Ladder has affected the morale of teachers. It used to be your morale was based on whether you felt you were a success. You could be the best teacher possible. Some teachers that are not nearly as good as other teachers have advanced or succeeded passing a certain stage and it's obvious that they are not as good a teacher as someone who did not, and that just shows it's not working. And it's going to make that teacher feel like 'Why should I give all I have anymore. What's the use? Maybe I should go and practice doing what it takes to pass the test and not worry about what goes on in my classroom.'"

Problems of distributive justice are significant not only because they may reduce the teaching commitment of the unrewarded; they may also inhibit school improvement if teachers cannot accept the legitimacy of conveyers' advice, assistance, and suggestions. How teacher selection for career ladders alters faculty interaction, then, has profound consequences for the ethos of the school.

Evaluation and collaboration. As a matter of fact, a serious and consistent foreboding by roughly a third of the teachers in our Southeastern sample about the evaluation instrument employed by the state was its threat to the positive collaborative relations teachers presently enjoyed. Teachers predict an end to offers of assistance among colleagues because all portfolio materials submitted by applicants for evaluation to the CLP had to be accompanied by evidence of their originality: "I'm not in favor of the Master Techer Plan. It's too much dog-eat-dog. I don't like the bit of someone getting an idea and wanting to close their door and not share anything with others. It hurts the children. When you do something good, you really ought to say, 'Hey, this really works well--you ought to try it'. Teachers are not going to do that if they have to document everything they've done as original. It's really hard because if you do something really well, you sure don't want anyone else to take credit for it. That's not the way education should be. That is not for the good of the children."

"One of the things that I'm worried might be negatively affected by the career ladder is sharing ideas. As part of my application I had to show them a portfolio of all the ideas I had accumulated. I think that we would wither and die if we couldn't share things with each other. I'm talking both about problems and a lot of good things that happen that we just want to share with others."

"I think the master teacher plan can lead to hoarding. We've talked about that in our school. It seems that if you don't desire master teacher status then you won't have the problem with hoarding. If you're really intent upon being a master teacher, then I think it might cause you to be a less sharing person."

"Teachers share a lot here. They are very professional. But the career ladder is changing all that. It used to be that we all had the same goals, so we helped each other. Now that recognition is being directed to individuals, everybody is trying to be their best to help themselves, not others."

"Teachers were more open and willing to share their ideas and their plans and work together and now it's kind of like 'Let me do my thing and make it as good as I can so I can make a good mark for me.' I really think it's detrimental."

If the means to select teachers for career ladder placement are ill-designed, if they are not thoroughly informed by an understanding of the nature of school effectiveness and the dynamics of group behavior, they are not likely to succeed. Should the costs of career ladders turn out to be the collegial relations needed to enhance teacher learning and commitment, in them we have the makings of a national educational failure at the very point in our history that we need a major success.

Quota systems. A separate issue in the design of career ladders is the question of scarce rewards--plans can either promote all who meet the standards of performance, or invite wide application and competitively choose the best qualified applicants from among those who apply (Murphy, Peterson & Kauchak 1985). To be sure, competitive rewards are lauded by some; Murphy et al. (1985) argue that teachers are more likely to accept a decision if they lose a promotion to another person rather than if the accuracy of the evaluation is at stake. Several states and localities do in fact create scarce rewards by limiting the number of teachers who can be recognized (e.g. Arizona, California, Utah).

But from a sociological point of view, we know that competitive rewards have unintended and negative consequences for group interaction. There is evidence that competitive rewards function to close rather than open communication and sharing among those who work together as well as to destroy trust. In competitive settings encouragement among group members is substantially reduced, and their problem-solving capacity diminished. In fact, competitive conditions may lead people to deliberately frustrate the attempt of others' to succeed (see Rosenholtz 1985b for a review).

Because skill development for teachers depends so heavily upon collaborative support and exchange, it seem reasonable to predict that competitive rewards will substantially thwart efforts to improve. Indeed, preliminary support for this assertion comes from a study of Great Britain's career ladder plan (Blomquist et al. 1984). In most British primary schools, where the range of differences in salary is modest, teachers share a closeness and work cooperatively. But in secondary schools, where salary ranges are far greater and where competition for promotions is keener, many teachers attempting to advance do not want to share their ideas unless they received full credit for them.

In fact, teachers involved with reforms studied by Blomquist et al (1984), Hart (1985), and Natriello & Cohn (1984) lowered their school commitment if they were not promoted. They became unwilling to perform tasks on the school's behalf unless there was personal benefit to be derived. And they resented those who were promoted, making school betterment an activity restricted, in all likelihood, solely to the chosen few.

The evolution of changes among teaching colleagues, therefore, becomes critical to document. How will collaborative exchange among teachers be affected by CLPs? What additional training will be needed to help master teachers succeed in their many functional assignments? What is the best mechnism for providing it? What are the characteristics of functional assignments that appear most promising in bringing about school improvement? How can competent teachers who are not selected for advancement still be made to feel appreciated? What will happen to their commitment to the school?

Clearly, much of the worth of policy research that I argue should commence, hinges on the fact that it will provide a reliable means of assessing the teaching occupation during a time of important changes, of monitoring those changes as they occur, and of supplying essential information, analysis and advice to those who will be making them occur.

Allocation inequities. The allocation of master teachers throughout a school system or region has been neglected by many CLPs, despite the fact that "good" schools have an easier time recruiting and retaining exemplary teachers than "bad" schools, and therefore have a disproportionate share (see Rosenholtz 1985a). That all schools need access to the valuable resource of good teaching seems obvious. Without small cadres of good teachers in every school, there is little support to ease transitions into teaching or to provide for the professional development of experienced or master teachers themselves.

The Charlotte-Mecklenberg Career Ladder Plan represents a significant departure from this omission, however. Here teachers advancing to the highest levels must be willing to transfer to different schools as need for their special skill arises. But what are the costs of high mobility? In Great Britain, allocative inequities are prevented by national advertisement and competition when positions arise. Due to the large number of promotions that are possible (8 different salary points along 5 scales), however, ambitious entrants to the career ladder, in order to advance rapidly, move frequently from one place to the next, developing little school commitment along the way (Blomquist et al. 1984). High mobility of this sort may have deleterious effects by limiting opportunity to develop the sort of collegial relations in schools that make teacher and student learning possible.

It is neither the intent of any CLP to make good teachers inaccessible to poorer schools, nor to encourage their mobility from schools before the impact of their efforts can be fully realized. These and other trade-offs that are inherent in the many structural changes proposed for the teacher workforce deserve and require research attention to enhance their understanding and impact.

CONCLUSION

The next decade will be a time of enormous turmoil in the teaching occupation. A majority of our teaching workforce in 1992 will be people who are not presently employed (NCES 1984). That means well over a million new teachers will be entering the classroom during the next six years. Who they are, how they will be trained and selected, what kinds of experiences and abilities they will bring with them, and what kinds of conditions they will encounter in the schools where they work are questions of more than academic interest. For this huge turnover is beginning just as the unsatisfactory quality of American schooling has seized the interest of policy makers at all levels to make changes intended to improve that quality. And the major object of these changes is the teaching workforce itself.

The combination of demographic forces and conscious policy decisions makes for a period of extraordinary volatility within and around the teaching force. There is also the eager anticipation--and hope--that through the many permutations of policy interventions, we will ultimately improve the current lackluster performance of schools. In reality, however, not enough is known about teachers and teaching to provide a steadfast base from which policy changes can be confidently launched. Where purposive efforts to improve quality are mounted, they may hit with highly uneven impacts if their effects are not properly anticipated. Further, without a satisfactory "feedback" mechanism, there is no avenue to supply continuing insight, constructive criticism, and dispassionate scrutiny to assist policy-makers in knowing whether their efforts are well-designed to solve actual problems or merely cosmetic changes that never penetrate beneath the surface.

The task of buttressing policy changes with real information, accurate analysis, and sound recommendation falls upon the research community. Such an ambitious enterprise has many dimensions: tracing and monitoring reform decisions; providing thoughtful and informed comment about them; offering technical advice to those who will be designing, implementing and evaluating them; and keeping in the public eye the conditions in education generally, and the teaching occupation particularly, that create compelling rationale for well-conceived changes. Only then can the promise of policy intervention become more than another episodic chapter in the history of American education.

References

Armor, D. J., Conry-Oseguera, P., Cox, M., King, N., McDonnell, L., Pascal, A., Pauly, E., & Zellman, G. (August, 1976). *Analysis of the school preferred reading program in selected Los Angeles minority schools*. Santa Monica, Calif.: The Rand Corporation.

Ashton, P. T., Webb, R. B., & Doda, N. (1983). *A study of teachers' sense of efficacy*. Final report, University of Florida.

Azumi, Jann E., and Serge Madhere. *Professionalism, Power and Performance: The Relationships between Administrative Control, Teacher Conformity, and Student Achievement*. Paper presented at the annual meeting of the American Educational Research Association, Montreal, April 1983.

Bishop, J. M. (1977). Organizational influences on the work orientations of elementary teachers. *Sociology of Work and Occupation*, *4*(2), 171-208.

Blomquist, R. *et al.* (1984). *Career Ladders in Great Britain: Implications for Utah Schools*. Report prepared for the Utah State Office of Education.

Bredeson, P. V., Fruth, M. J., & Kasten, K. L. (1983). Organizational incentives and secondary school teaching. *Journal of Research and Development in Education*, *16*, 52-56.

Bridges, Edwin M., and Maureen T. Hallinan. Subunit Size, Work System Interdependence, and Employee Absenteeism. *Educational Administration Quarterly* 14, no. 2 (1978): 24-42.

Brookover, W., Beady, C., Flood, P., Schweitzer, J., & Wisenbaker, J. (1979). *School systems and student achievement: Schools can make a difference*. New York: Praeger.

Bruno, James E. (1981b). Morale-affecting Stressors: An Analysis of Black, White, and Hispanic Elementary Schools. *Urban Education* *16*.

Bruno, J. E. (1981a). Designs of incentive systems for staffing racially isolated schools in large urban districts: Analysis of pecuniary and nonpecuniary benefits, *Journal of Educational Finance*, *7*.

Bruno, James E., and Mary Lynn Doscher. Contributing to the Harms of Racial Isolation: Analysis of Requests for Teacher Transfer in a Large Urban School District. *Educational Administration Quarterly* 17, no. 2 (1981): 93-108.

Burlingame, Martin, *An Analysis of Attrition in the Downstate Illinois Teaching Force, 1972-73 to 1977-78*. Springfield, Illinois State Board of Education, 1980.

Campbell, J. P. & Pritchard, R. D. (1976). Motivation theory in industrial and organizational psychology. In M. Dunnett (ed.) Handbook of Industrial and Organizational Psychology, New York: Rand-McNally.

Chapman, D. W. (1984). Teacher retention: the test of a model. American Educational Research Journal, 21(3), 645-659.

Chapman, D. W., & Hutcheson, S. M. (1982). Attrition from teaching careers: A discriminant analysis. American Educational Research Journal, 19, 93-106.

Chapman, D. W., & Lowther, M. A. (1982). Teachers' satisfaction with teaching. Journal of Educational Research, 75(4), 240-247.

Charters, W. W., Jr. (1970). Some factors affecting teacher survival in school districts. American Educational Research Journal, 7, 1-22.

Clark, D. L., Lotto, L. S. & Astuto, T. A. (1984). Effective schools and school improvement: A comparative analysis of two lines of inquiry. Educational Administrative Quarterly, 20, 41-68.

Cohen, E. G. (1973). Open-spaced schools: The opportunity to become ambitious. Sociology of Education, 46, 143-161.

Cook, K. and Hegtvedt, K. A. (1983). Distributive justice, equity, and equality. Annual Review of Sociology, 9, 217-241.

Cox, Harold, and James R. Wood. Organizational Structure and Professional Alientation: The Case of Public School Teachers. Peabody Journal of Education 58, no. 1 (1980): 1-6.

Cruikshank, D. R. (1985). Uses and benefits of reflective teaching. Phi Delta Kappan, 66, 704-706.

Darling-Hammond, L. & Wise, A. E. (1985). Beyond standardization: State standards and school improvement. Elementary School Journal, 85, 315-336.

Frataccia, E. V. & Hennington, I. (1982). Satisfaction of hygiene and motivation needs of teachers who resign from teaching. Paper presented at the annual meeting of the Southwest Research Association, Austin.

Gehrke, N. J., & Kay, R. S. (1984). The socialization of beginning teachers through mentor-protege relationships. Journal of Teacher Education, 35, 21-24.

Glenn, B. C., & McLean, T. (1981). What works? An examination of effective schools for poor black children. Cambridge, Mass.: Harvard University, Center for Law and Education.

Glidewell, J. C., Tucker, S., Todt, M., & Cox, S. (1983). Professional support systems: The teaching profession. In A. Madler, J. D. Fisher, & B. M. DePaulo (Eds.), Applied research in help-seeking and reactions to aid. New York: Academic Press.

Hart, A. W. (1985). *Case study prepared for the Utah State Office of Education*. Unpublished manuscript, University of Utah.

Hawley, W. & Rosenholtz, S. J. (1984). Good schools: What research says about improving student achievement. *Peabody Journal of Education*, (61), 1-178.

Hort, S., Stiegelbauer, S. & Hall, G. (1984). How principals work with other change facilitators. *Education and Urban Society*.

Hoy, W. K. Pupil Control Ideology and Organizational Socialization: A Further Examination of the Influence of Experience on the Beginning Teacher. *School Review 77* (1969): 257-265.

Hoy, W. K., C. J. Tarter, and P. Forsyth. Administrative Behavior and Subordinate Loyalty: An Empirical Assessment. *Journal of Educational Administration* 16, no. 1 (1978):29-38.

Huberman, A. Michael and Miles, M. (1984). Rethinking the quest for school improvement: some findings from the DESSI study. *Teachers College Record*, *86*, 34-54.

Johns, G., and Nicholson, N. (1982). Meanings of absence: new strategies for theory and research in B. M. Staw and L. L. Cummings eds. *Research in Organizational Behavior*, *4*.

Katzman, M. T. *The Political Economy of Urban Schools*. Cambridge, Mass: Harvard University Press, 1971.

Levin, H. M. A Cost-effective analysis of teacher selection. *Journal of Human Resources*, *1*.

Litt, M. D., & Turk, D. C. (April, 1983). *Stress, dissatisfaction, and intention to leave teaching in experienced public high school teachers.* Paper presented at the annual meeting of the American Educational Research Association, Montreal.

Little, J. W. (1982). Norms of collegiality and experimentation: Workplace conditions of school success. *American Educational Research Journal*, *19*(3), 325-340.

Locke, E. A. (1976). The nature and causes of job satisfaction. In M. Dunnett (ed.) *Handbook of Industrial and Organizational Psychology*. New York: Rand-McNally.

Lortie, D. C. (1975). *Schoolteacher: A sociological study*. Chicago: University of Chicago Press.

Mann, D. & Fenwick, S. (1985). *The Impact of Impact II: 1982-1983.* Report to the Exxon Education Foundation, Teachers College, Columbia University.

March, J. G., & Simon, H. A. (1958). *Organizations*. New York: John Wiley & Sons.

McArthur, John T. What Does Teaching Do to Teachers? Educational Administration Quarterly 14, no. 3 (1978):89-103.

McLaughlin, M. W., & Marsh, D. D. (1978). Staff development and school change. Teachers College Record, 80(1), 69-94.

Murphy, J. & Hallinger, P. (1982). The role of the superintendent as instructional leader. Administrator's Notebook, 31.

Murphy, M. J., Peterson, K. D., & Kauchak, D. P. (1985). Analysis of Utah Career Ladder Plans. Report prepared for the Utah State Office of Education.

National Center for Education Statistics (1984). The Condition of Education, Washington, D.C.

Natriello, G. (1984). Teachers' perceptions of the frequency of evaluation and assessments of their effort and effectiveness. American Educational Research Journal, 3, 579-596.

Natriello, G. & Cohn, M. (1985). Critical issues in the development of a merit pay system. Administrator's Notebook, (31).

Natriello, G., & Dornbusch, S. M. (1980-81). Pitfalls in the evaluation of teachers by principals. Administrator's Notebook, 29(6).

Pavalko, R. M. (1970). Recruitment to teaching: patterns of selection and retention, Sociology of Education, 43.

Pederson, K. George. "Teacher Migration and Attrition." Administrator's Notebook 18, no. 8 (1970):1-13.

Phi Delta Kappa. Why Do Some Urban Schools Succeed? The Phi Delta Kappa Study of Exceptional Urban Schools. Bloomington, Ind.: Phi Delta Kappa, 1980.

Popham, I. J., Curse, K. L., Rankin, S. C., Sandifer, P. D., and Williams, P. L. (1985). Measurement-driven instruction: it's on the road. Phi Delta Kappan, 66, 628-634.

Raschke, D. B., Dedrick, C. V., Strathe, M. I., and Hawkes, R. R. Teacher stress: the elementary teacher's perspective, Elementary School Journal, 85, 559-564.

Rosenholtz, S. J. (1985a). Effective schools: interpreting the evidence. American Journal f Education, 93, 352-388.

Rosenholtz, S. J. (1985b). Political myths about education reform: lessons from research on teaching. Phi Delta Kappan, 66, 349-355.

Rosenholtz, S. J., Bassler, D. & Hoover-Dempsey, K. (1985). Organizational inducements for teaching. Interrim report submitted to the National Institute of Education, University of Illinois.

Rosenholtz, S. J. & Greer, E. (1985). *Learning to teach: experience or setting?* Unpublished manuscript, University of Illinois.

Rutter, M., Maughan, B., Mortimore, P., & Ouston, J. (1979). *Fifteen thousand hours: Secondary schools and their effects on children.* Cambridge, Mass.: Harvard University Press.

Schlechty, P. C., & Vance, V. S. (1983). Recruitment, selection and retention: The shape of the teaching force. *Elementary School Journal*, *83*, 469-487.

Sergiovanni, Thomas. Factors Which Affect Satisfaction and Dissatisfaction of Teachers. In *The Employment of Teachers*, edited by D. Gerwin. Berkeley, Calif.: McCutchan Publishing Corp., 1974.

Sherwin, J. S. (1969). *Four problems in teaching English: A critique of research.* Scranton, Penn.: International Textbook.

Sizemore, B., Brossard, C. A., Harrigan, B. (1983). *An abashing anomaly: The high achieving predominantly black elementary school* (Final report). Pittsburg: University of Pittsburg.

Spuck, D. W. (1974). Reward structures in the public high school. *Educational Administration Quarterly*, *10*, 18-34.

Suhor, C. (1985). Objective tests and writing samples: how do they affect instruction in composition. *Phi Delta Kappan*, *66*, 635-639.

Summers, A., & Wolfe, B. (1977). Do schools make a difference? *American Economic Review*, *67*, 639-52.

Turnball, B. J. (1985). Using governance and support systems to advance school improvement. *Elementary School Journal*, *85*, 337-352.

Venezky, R. L., & Winfield, L. F. (1979). *Schools that succeed beyond expectations in reading. Studies in education.* Newark, Del.: University of Delaware.

Warren, Richard L. Context and Isolation: The Teacing Experience in an Elementary School. *Human Organization* 35 (1975): 139-48.

Weaver, W. T. (1981). Demography, quality and decline: the challenge for schools of education for the 1980's in *Policy for the Education of Educators: Issues and Implications.* Washington, D.C.: American Association for Colleges of Teacher Education.

Willower, Donald J., and Ronald G. Jones. When Pupil Control Becomes an Institutional Theme. *Phi Delta Kappan* 45 (1963): 107-09.

Assessing American Education:
Shrinking Resources, Growing Demands

Diane Scott-Jones
Department of Psychology
North Carolina State University

Paper commissioned by the
National Center for Education Statistics
June, 1985

Assessing American Education: Shrinking Resources, Growing Demands

National statistics on education are amassed in order to monitor the extent to which individuals are being educated and the quality of education they receive. These statistics are necessary for planning future educational efforts and for assessing the impact of past and current educational policy and practice. Few people appear satisfied with the American educational system. In 1983, the "year of the reports", several studies critical of American education were published (Howe, 1984). These reports (e.g., National Commission on Excellence in Education, 1983) call for greater attention to "excellence" in education. At the same time, many individuals insist that the race/ethnicity, class, and gender biases pervasive in American society must not be reflected in children's school experiences or in their achievement levels (see, for example, Harvey, 1985; S. Klein, 1985). Further, schools are expected to do more than focus exclusively on the teaching of narrowly defined cognitive skills. Because it is an almost universal experience for American children, occupying much of their lives, schooling is expected to facilitate the healthy development of children into competent, well-adjusted adults. In short, many varied demands are being made of the educational system, at a time when public support, particularly at the federal level, is declining.

In a time of shrinking funds for education, one could argue that scarce resources should be channelled into direct services for children. Careful monitoring of education, however, becomes even more critical in times of financial difficulty so that the effects of cutbacks and of continued spending can be documented. Bell (1982) includes conducting and financing educational research, and strengthening national research capabilities, especially at universities, as legitimate functions in a limited federal role in education. Even a conservative position, then, acknowledges the strong need for educational research. Data collection should be carefully planned, to avoid unnecessary expenditures and repetition. Collaboration among agencies and individual researchers will be necessary. Efforts such as the recent Interagency Conference on Child and Family Statistics (Zill, Peterson, & Moore, 1984), aimed at coordinating and improving national statistics on children generally, including those on education, should be continued. Although funding for data collection is a serious issue, the cost of various data collection programs will not be addressed directly in this paper. Wisdom and efficiency in spending are important goals; however, it is a truism that many other areas of governmental spending are far more wasteful and less useful to the society than is the education of America's children.

This paper describes critical issues in elementary and secondary education and the manner in which national data collection efforts might address these issues effectively. These issues include, first, the equal importance of equity and excellence as goals of American education. Several phenomena related to the goals of equity and excellence are discussed, including bilingual education, private schools, use of computers in schools, drop-outs, and the transition from school to work. A second major issue is the development of students from childhood through adolescence. Developmental issues include the need for longitudinal studies, especially of elementary school children, preprimary school programs, and adolescent development. Finally, the quality of instruction, including teacher preparation and classroom processes, and public perceptions of education are discussed.

Excellence and equity in education

A foremost goal of the American educational system is the development of students into literate, critical-thinking citizens who function well in a complex, technological, democratic society. The elementary and secondary educational system should prepare students for adult roles that are both productive and personally satisfying. For some students, this preparation will lead to higher education and for others, the transition into the labor force.

The 1983 reports expressed the fear that American children's education is not sufficient to allow the country to compete favorably with other industrial nations. Appropriate indicators of the quality of education American children receive should be available and should allow for meaningful comparisons among states, as well as with other countries. In 1980, 40 states had minimum competency testing (Whalen, 1984) but these programs varied greatly from state to state. Indicators such as SAT and ACT scores, reported in the State Education Statistics Chart, have some utility but do not allow complete across-state comparisons, because the percentage of students taking each test varies from state to state. An exclusive reliance on standardized, multiple-choice test formats, however, is not desirable. Measures of critical thinking must be available. Test items are needed that assess higher-level processes rather than only the recall of basic facts. Students' ability to write must be assessed. Other abilities and skills judged to be important in the society should be assessed. For example, do we expect our students to have some knowledge of music, art, languages other than English? Is physical and nutritional education an important issue, given the recent reports of low physical fitness in American school children? If so, these must be assessed in some meaningful way.

Equity issues. The SAT and ACT also have limited utility as indicators of the status of American education because segments of the population will not take these tests and will not attend college. This fact represents a most difficult issue. Excellence and equity must be simultaneous goals of the educational system; both must be included in assessments of educational outcomes. "Quality" and "excellence", however, appear to have become codewords for lessened concern with equality and equity in education. Some proponents of excellence in education appear to believe that the only way to achieve that goal is by focusing on those children already best-served by the American education system and making them better.. Other educators claim to believe that equity for minorities is no longer an issue, except in the most minor ways, because of the changes of the 1960s and 1970s. In outlining a limited federal role in education, Bell (1982) includes protection against discrimination and violations of civil rights, but with an emphasis on persuasion rather than enforcement. This statement exemplifies the current lessened interest in equality of opportunity and outcomes for those groups in society underserved by the educational system. Excellence in education, however, should be a goal for all students. Quality without equality will mean, simply, continued discrimination.

Equity for the poor, minorities, and females is a major theme within all the issues raised in this paper. (Equal access to education for handicapped children also is an important issue, but will not be addressed directly.) A growing number of American children live in poverty. Currently, 22% of American chidren live in poverty, compared to 18.3% in 1980. Poverty levels vary greatly by region and state, ranging from 7.5% of the school-age population in Wyoming to 30.4% in Mississippi. Minorities account for 26.7% of public school enrollment, ranging dramatically from .9% in Maine to 96.4% in the District of Columbia (U.S. Department of Education, 1985). Although they may be approximately equal in number with their male counterparts, girls, along with the poor and minorities, continue to experience inequities in the educational system. The achievement of poor, minority, and female students thus is a major issue.

Many reports do not provide adequate information on race, sex, and income level of students. Reports of education statistics provide some separate data for Blacks but not always in the most informative manner. Little information about Hispanics is provided and even less about other minority groups, such as American Indians. (See La Fromboise & Plake, 1983, for discussion of research needs of American Indians). Data sometimes are reported by race and then separately by sex, rather than by sex within race. Gender differences, however, may not be uniform for whites and minorities (Reid, 1982; Scott-Jones & Nelson-Le Gall, in press). Similarly, breakdowns by economic status are given for the entire sample rather than within minority groups. Data should be reported by income level within minority groups, to avoid the usual confounding of race and economic status. In those instances where some information on income level is provided, it is usually at the aggregate rather than the individual level. For example, reports of mathematics performance from the National Assessment of Educational Progress (NAEP) (Weinberg, Gerald, & Tron, 1984) give mean performance by type of community--rural, urban disadvantaged, and urban advantaged.

The achievement of poor, minority, and female students must be carefully assessed. The effective schools movement has directed attention to achievement in schools with large numbers of poor minority students and has highlighted the fact that poor minority children can benefit from good educational efforts. It has focused attention, however, on standardized test scores as the major criterion for effective schooling. Assessment should include comprehensive measures of learning and thinking that are appropriate for minority students. Sex differences in performance, especially in math and science (see Stage, Kreinberg, & Eccles, 1985), need to be monitored.

Integration of public and private schools must be monitored. In addition, segregation of students via ability grouping and tracking must be monitored. Ability grouping in classrooms may result in an inferior education for minority and poor students. Minority students tend to be overrepresented in lower level groups and tracks, where teachers spend less time on academic instruction and more time on discipline and classroom management (Hallinan, 1982).

A biennial or triennial report on the education of children living in poverty was suggested at the Interagency Conference on Child and Family Statistics (Zill, Peterson, & Moore, 1984). Reports on the education of racial/ethnic minorities also would be useful. Oversampling of minority groups was recommended by the Interagency Conference on Child and Family Statistics. Because national probability samples will include only small numbers of Blacks, Hispanics, Native Americans, and Asian Americans, oversampling of these groups may be necessary for meaningful conclusions about them. Supplementary samples could be obtained from geographic areas with high numbers of specific minority groups. The whole issue of cultural pluralism needs to be addressed. Not only are numbers of minorities in schools important but also respect for and acknowledgment of people of different racial and ethnic backgrounds.

Bilingual education. Bilingualism is an issue in some regions. Children may be labeled "language minority" if they or their parents use a language other than English primarily or "often". Students are designated "limited English proficient" if they score below a cut-off point on a test of English proficiency (Weinberg, Gerald, & Tron, 1984). States with the highest percentages of low-scoring children are New Mexico, New York, California, Arizona, Hawaii, and Texas. Approximately 73% of the 3.6 million children labeled limited English proficient are Hispanic (Rotberg, 1982), although the majority of U. S. Hispanic children do not attend bilingual classes (Otheguy, 1982). Because of their greater numbers, Hispanics have been the subject of research more frequently than other language minority groups, such as Asian Americans or American Indians (Steinberg, Blinde, & Chan, 1984), although the court case, Lau v. Nichols, 1974, resulting in the establishment of bilingual education, was brought on behalf of Chinese students (Sinclair, 1983).

Disagreement exists regarding the number of children who need special English language services (Cooke, Ginsberg, & Smith, 1985; NCES Indicators, 1985). Cooke et al. point out that many children scoring below the cut-off use English as their main or only language and therefore do not need special English services. Unless the test is not valid or the cut-off point is too high, however, it seems important to provide services to low-scoring children, even if they do use English. Currently available tests of language proficiency are of questionable reliability and validity, and use norms derived from nonrepresentative samples (Padilla & Lindholm, 1984; Rotberg, 1982). Most important, the tests do not measure the functional use of language (Padilla & Lindholm, 1984). Determining language proficiency for bilingual children is difficult because fluency in a language is dependent upon many factors such as the social context, the language permitted or encouraged in that context, and the topic of conversation (Duran, 1984). For example, Hispanics may use Spanish generally but may be more familiar with English as the main language of the school (Duran, 1984).

In addition to the identification of children needing special English services, another issue is the nature of the special instruction they receive. Researchers disagree on whether the focus of special programs should be immersion in the English language in special classes, with instruction in other subjects in the usual classes, or bilingual programs that teach children in their native language. Clear differences in student outcomes in the two types of programs have not been found. Although methodological problems preclude definitive conclusions, some studies suggest that bilingual programs result in lower drop-out rates (reported to be as high as 90% for Hispanics in a Texas school district), better attendance, and higher self-concepts (Otheguy, 1982; Rotberg, 1982). In studying the effects of different programs, other variables such as socioeconomic status, ethnic or national origin, and age on arrival in the U.S. should be included (Rotberg, 1982; Steinberg et al., 1984). In addition to achievement as a major outcome of bilingual programs, the preservation of the child's native language may be a desired goal for many, although the issue is controversial (see Orteguy, 1982). Spanish-English bilingual children in Head Start bilingual programs appear to decline in complex Spanish linguistic forms as they increase in complex English forms (Garcia & Gonzalez, 1984).

Identifying qualified teachers and appropriate curricula may be a problem. For example, because of the Title VII (Elementary and Secondary Education Act, 1974) ruling that 20 or more students from the same language group in a given school district must get special instruction, the Chicago school district must provide instruction in Spanish and 17 other languages. A study in New Mexico found that only 13 of 136 bilingual teachers and aides could read and write Spanish at the third grade level (Rotberg, 1982).

Segregation becomes an issue in the education of bilingual children. To avoid segregation, Title VII funding of bilingual programs allows up to 40% of the children enrolled to be native

English-speakers (Rotberg, 1982). Hispanic students may be assigned to bilingual programs on the basis of ethnicity or home language rather than on language proficiency. In a study of Title VII bilingual programs, 75% of the students were Hispanic but less than one-third of them were judged by a teacher to be of limited English proficiency (Rotberg, 1982).

Private schools. Many parents view public schools, long the backbone of American education, as less desirable than private schools. The expectation of higher achievement and of a relatively homogeneous student body may lead some parents to prefer private schools. A report comparing public and private schools, prepared for NCES by Coleman, Hoffer, & Kilgore (1981), generated much controversy regarding the segregation of private schools, the differential achievement of private and public school students, and the predicted effects of public support of private schools. Critics of the report (e.g., Braddock, 1981a, 1981b) point out that the finding of little segregation in private schools is meaningless, because of the miniscule number of minority students in private schools. The index of segregation used in the report was based on the distribution of Black and white students within the private school system, not on the proportion of Blacks in private schools. It assessed whether those Blacks within the system of private schools tend to be segregated in certain schools. The important issue, however, is Black access to private schools, not their segregation within the private school system.

The finding of higher achievement for private schools does not take into account self-selection artifacts or curriculum placement of public and private school students. The higher achievement found for private school students might be diminished if they were compared to public school students in college preparatory tracks rather than the entire public school population (Braddock, 1981a, 1981b).

Finally, the Coleman et al. claim that public support of private schools would aid poor and minority students and would not lead to further segregation is not supported (Braddock, 1981a, 1981b). The support and improvement of public schools are necessary for educational opportunity for poor and minority students. Public and private school comparisons made by Coleman et al. yield more useful information about the characteristics of effective schools--such as emphasis on homework, demanding curricula, demanding teachers, and a disciplined, orderly environment--which could be used to improve public schools (Braddock, 1981a, 1981b).

The biennial surveys of private schools should be continued and strengthened. Efforts must be made to ensure that the samples in private school surveys are representative and that reports from administrators of these schools are reliable. Because of the lax regulation of private schools, data regarding enrollment, attendance, achievement, and teacher certification may be suspect. Regulations for private schools vary by state.

In North Carolina, for example, private schools are required to report only the beginning and permanent termination of operation (North Carolina General Statutes, 1979). They are not required to hire certified teachers. Participation in state achievement testing and minimum competency testing programs is voluntary, although private schools are required to give a standardized achievement test to students at specified grade levels and a standardized minimum competency test to eleventh graders each year. They must maintain records of these tests, along with attendance and immunization records, for one year for possible inspection; the appropriate state agency may or may not inspect these data, at its own discretion. The high-achieving private schools described in the Coleman et al. study may not be the norm.

The stability of private schools--how long they stay in operation-- needs to be monitored, along with patterns of movement of students between private and public schools. Individual level data from parents and students should be obtained, particularly regarding motives for attending private schools and perceived and actual benefits of private school attendance. The motives and experiences of Blacks and other minorities in private schools should be assessed, as has been done in Slaughter's recent work (Slaughter & Schneider, 1985; Slaughter, Schneider, & Lindsey, 1985).

Some parents eschew formal schooling altogether, choosing instead to teach their children in their homes. In North Carolina, a recent court ruling required that these parents comply only with the minimal regulations for private schools; they are not required to report anything other than the beginning and the permanent closing of operation (Alvarado, 1985). In 1985, at least twenty-five "schools" with students from two or three families were established in North Carolina. Officials expect the number of home schools to increase. A parent group lobbying for home education in North Carolina claims 300 families as members (Perkins, 1985). More information is needed on the numbers, experiences, and achievement of children taught in their own homes.

Computers and schools. Changes in technology in the society affect schools. An increasingly computerized society requires that children become computer functional, if not computer literate. The National Commission on Excellence in Education (1983) recommended that computing be a basic high school subject. The extensive use of computers may result in changes in definitions of intelligence and education: education may move from basic facts to information management skills; broad problem-solving skills; planning, monitoring, learning to learn, and other metacognitive skills; and communication and inquiry skills (Pea, 1985). Research on the computers' impact is especially needed now, before computers become so widespread that finding children unfamiliar with computers will be difficult (Lepper, 1985).

NCES has assessed the availability of microcomputers and of terminals to mainframes in elementary, junior high, and senior high schools (Grant & Snyder, 1983). From 1982 to 1983, the percentage of elementary schools having microcomputers more than tripled, the percentage of junior high schools more than doubled, and the percentage of high schools tripled (Whalen, 1984). The amount of time students use computers was assessed in a study conducted by Johns Hopkins' Center for the Social Organization of Schools. In elementary schools, the median number of minutes of student use per week was low: 12 minutes for learning and recreational games, 13 minutes for drills and remedial work, and 19 minutes for programming and computer literacy; almost no students spent more than 60 minutes per week using the computer. The biggest increase from elementary to secondary schools was in programming and computer literacy (median, 55 minutes) and in word processing and data processing (30 minutes). These amounts are not adequate. At least 500 hours of computer time may be necessary for the development of expert programming skills (Linn, 1985).

Although these surveys provide useful beginning information about computers in schools, other issues must be addressed. Computer technology, in its present transitional stage, appears to amplify existing social roles and social problems (Caporael & Thorngate, 1984). The use of computers may increase the achievement gaps between middle-class children, who are likely to have and use computers in their homes as well as in school, and poor children, who may have only limited access to an inadequate number of computers at school. According to Linn (1985), access to computers at home and at school is related to middle-school students' programming skills, except in those classes labeled "exemplary". In addition, boys appear to master and enjoy computers more than do girls. Boys' greater affinity for computers may result from the preponderance of educational programs with male sex-typed themes such as sports, war, and violence (Lepper, 1985) and from the association of computers with mathematics in junior and senior high schools (Sheingold, Kane, & Entreweit, 1983). When girls take computer courses, they may perform as well as boys. In Linn's (1985) research, girls were 37% of middle school programming classes but 60% of students identified as the most talented programmers. These trends that could maintain or exacerbate existing inequities need to be monitored. Computer-assisted instruction may be useful for poor minority students in some situations, however. Sheingold et al. (1983) describe teen-aged Asian immigrants, attending school for the first time, who had successful experiences learning mathematics via computers, because little English was required.

Attention must be given to the actual use of computers and what students gain from them. Derek Bok of Harvard University has suggested that computers, because they limit the student to a specified set of responses, may restrict students' imagination (Culliton, 1985). According to Bok, some areas of study, such as historical interpretation or literary criticism, cannot be reduced to the formal rules and procedures necessary for computerized instruction. Ideally, students should not be passive, using the computer only for routine drills; such students may guess randomly just to move the program along.

Although computer drills may not be much of an advancement over workbook drills, other uses of the computer may facilitate learning in unique ways. Carefully designed tutorial programs may allow for the redirection of students' learning according to the nature of their errors. Word processors may encourage the development of skills with written language, because of the ease of editing and revising (Wolf, 1985); computer simulations of principles and experiments in the physical and biological sciences may help to make abstract concepts concrete and understandable (Chaille & Littman, 1985); computer graphics may enhance the development of spatial skills (Klein, 1985). Learning may be facilitated most when children acquire the control that results from learning to program the computer (Sheingold, et al., 1983). Children who learn to program may acquire valuable conceptual and problem-solving skills that will be useful in other contexts (Dickson, 1985; Lepper, 1985; Linn, 1985; Olson, 1985).

Research is needed to determine whether computer-assisted instruction actually enhances learning and achievement. Research should address whether computer-assisted instruction in a "fun-and-games" format is more or less effective than typical "drill-and-practice" computerized presentations (Lepper, 1985). Research should address whether programming skills actually transfer to other problem-solving situations. Attention should be given to the possibility that the computer may maintain achievement gaps between low achievers who use them for drill and high achievers who learn to program.

Computers may change the social dynamics of the classroom. Hawkins, Sheingold, Gearhart, & Berger (1982) found that elementary school teachers focus on the social rather than the cognitive outcomes of using computers in their classrooms. Contrary to what might be expected, more social interaction, collaboration, and helping are observed among students working on computers than among those working on other tasks. Dickson (1985) suggests that the potential for increased social interaction in classrooms with computers, expecially those with "thought-provoking" software, may be more significant than the potential for increased cognitive skills. Social interaction related to cognitive tasks contributes to cognitive development (Chaille & Littman, 1985).

The appropriateness of software and its relationship to the curriculum should be monitored. Software may be independent of the curriculum, it may be chosen to reflect the existing curriculum, or it may result in the design of a revised curriculum incorporating the computer. Bok suggests that the process of software development, because it requires detailed attention to the presentation of material in a manner that facilitates student learning and interest, may lead to a more careful appraisal of the entire instructional process (Culliton, 1985). The use of computers may also lead to a more individualized curriculum (Lepper, 1985; Sheingold et al., 1983).

Teacher expertise in using computers needs to be monitored. Teachers may need time to develop their skills in the school rather than more formal courses (Sheingold et al., 1985). Some teachers become "computer buffs" on their own initiative but the active support of the principal in acquiring time and resources may be critical (Sheingold et al., 1985). As computers become integrated into curricula, aides may be able to manage students' computer activities and teachers may focus on higher-level conceptual development and social development (Sheingold et al., 1985). The emergence of student computer experts in some schools has led to new student-teacher relations, with students sometimes contributing to the development of teachers' skills (Sheingold et al., 1985). Teachers interacting with student-experts may experience role conflict (Caporael & Thorngate, 1984) but this may be a temporary phenomenon that will disappear as computers become more widely used and the usual structure of expertise resumes (Sheingold, Hawkins, & Char, 1984).

Drop-out rates. The rate at which children leave school without a high school diploma needs to be monitored. The drop-out rate, or its inverse, the graduation rate, is considered an important indicator of the status of American education. A major performance outcome in across-state comparisons, the graduation rate varies from a reported high of 94.8% in North Dakota to a low of 57.2% in Louisiana (U.S. Department of Education, 1985).

There are discrepancies between the drop-out rates estimated by NCES (27%) and by the Census Bureau (16%) (Cooke et al., 1985). The NCES measure is the difference between the number of public high school graduates in a given year and the number of public school 9th-graders from four years before; the Census Bureau, in household surveys, asks individuals how many years of school they have completed. Cooke et al. point out a number of reasons for the possibly inflated NCES measure and the possibly too-low Census figure. The basic problem, however, is one of definition. The Census measure is actually an assessment of the proportion of the population reporting that they have completed high school, including those who pass high school equivalency exams. These self-report data are not checked against any official data. The NCES statistic, if it excluded current graduates who were not ninth-graders four years earlier, would be a clean measure of the proportion of ninth-graders who graduate on schedule. Follow-up assessments could determine the rate at which the ninth-graders graduate five or six years later. This assessment still would miss students who left school before ninth grade.

The High School and Beyond study provides useful information about drop-outs from the 1980 sophomore class (Whalen, 1984). Data are reported by sex, race/ethnicity, socioeconomic status, grade and curriculum, community type and geographic region, and public or private school. Most important, the reasons students give for dropping out are reported by sex and race. Reasons cited most frequently were dislike of school for white males, poor grades for minority males, marriage and dislike of school for white females, and poor grades and pregnancy for minority females. This kind of information is critical for the prevention of students' dropping out. In addition, distinctions should be made between truancy, chronic absenteeism, and drop-outs (Zill et al., 1984). Some students who have not formally left school may be absent so frequently that they, for all practical purposes, are drop-outs (Steinberg et al., 1984). More information on drop-outs in minority groups is needed. Drop-out rates for Hispanics (40%) and American Indians (38-60%) are extremely high, much greater than those of Blacks and whites (Steinberg et al., 1984).

Transition to work. Much attention is given to the preparation of students for higher education, as evidenced by the use of the SAT and ACT in comparisons of educational outcomes across states (U.S. Department of Education, 1985). Many students, however, want and need to enter the job market immediately after high school. They need to be prepared for an occupation and for forms of additional training other than college. The transition directly from high school to work occurs more often for poor and minority students than for middle-class and white students.

A most serious dilemma is the relationship of the education children receive to the nature of employment available to them in the society. Current evaluations of American education, such as that of the National Commission on Excellence in Education (1983), suggest that the faltering American economy, both its domestic and international status, is the result of our poor educational system, or at least can be revived by improvements in the educational system. Giroux (1984) points out that the converse is true--the economy has a great impact on the schools. Duckworth (1984) cites Bureau of Labor Statistics projections of more fast-food than high-technology jobs in 1990. The increasing polarization of available jobs--a relatively small number of high technology jobs and a high number of low-level service jobs requiring few intellectual skills--raises difficult questions about the economy, the educational system, and the complex relationship between the two (Giroux, 1984).

Although some may argue that a successful American economy depends upon well-educated, highly skilled adults, such may not be the case. If projections regarding available jobs are accurate, the economy may need large numbers of adults willing to work in what are presently low-paying, unsatisfying service jobs. Are we trying to teach children skills they will not be able to use as adults? If the majority of children acquire high-level technical and conceptual skills, how will the society decide who gets the more lucrative, prestigious high-tech jobs and who gets the service jobs? Does the achievement of poor minorities remain low because of perceptions of a "job ceiling" (Ogbu, 1978)?

These questions regarding the utility and promise of the educational system for all of America's youth must be answered. A continuation and strengthening of the assessment of students' activities after high school is necessary. The nature of employment, as well as the rate of employment, should be assessed.

Childhood development

Longitudingal studies. Because schools provide a major context for children's development, much more information is needed on children's experiences in school, in addition to the outcomes of schooling. A national longitudinal survey of children in the elementary grades, similar to the longitudinal surveys focusing on secondary schools, has been proposed (Zill et al., 1984). Such a survey should include children and parents as respondents in addition to gathering information about classroom processes. Students' attitudes toward school and achievement aspirations and expectations should be assessed. The relative inattention to elementary school children in national longitudinal studies is not in keeping with the importance of development during this time of children's lives.

Preprimary schooling. Because of dramatic changes in family structures, one cannot assume that most school children live in homes with a father who is the sole breadwinner and a mother who is a full-time homemaker. More than half of women with children work outside the home and a high proportion of children live with a single parent, usually the mother, for at least a portion of their lives. Minority children are more likely than white children to live in single-parent homes. These changing family structures have implications for after-school programs, and for day care services as part of public schools. The proportion of 3- and 4-year-olds enrolled in some form of schooling almost quadrupled between 1964 and 1983, rising from 9.5% to 37.5% (NCES Indicators, 1985). The educational experiences of 3- and 4-year-old children will continue to be an important issue. The survey of prekindergarten enrollment, conducted through the Census' Current Population Survey, should be continued. A related issue is the before- and after-school care of children of working parents, which often is handled by the same providers who take care of 3- and 4-year-olds. Lack of appropriate care has resulted in the problem of "latch-key" children.

Although day care facilities that have been studied appear to cause neither increases nor decreases in educational achievement for children generally (Belsky & Steinberg, 1978), longitudinal research shows positive educational and economic outcomes for children, mostly poor and minority, who participate in specially designed prekindergarten programs (Lazar & Darlington, 1982). The quality of present day care facilities is an issue of concern; some facilities are exemplary but others are merely custodial and some have allowed the abuse of children enrolled. The cost as well as the availability of quality care is an issue; many parents, especially single mothers, simply cannot afford good day care.

Some European and Asian countries presently provide educational programs supported by public funds for 3- and 4-year-olds. In the United States, some states have public programs, which typically are half-day programs and are not available to all 3- and 4-year-old children; other states are planning programs for 1985-86 and some are studying the need for and feasibility of such programs. North Carolina, for example, recently completed a study of the issue (Kahdy, 1985). They concluded that the public school does have a responsibility to provide programs for prekindergarten children because of the developmental needs of children of that age group and because of the needs of working parents. The most vigorous objection to public prekindergarten programs came from private day-care providers, with some concern also expressed by those who feared that structured programs and pressure for academic success would overburden children.

In analyzing and reporting prekindergarten enrollment data, emphasis should be given to differential enrollment rates of minority groups. In 1981, approximately 36% of both Black and white 3- and 4-year-olds were enrolled in school programs, compared to approximately 25% of Hispanic children (Grant and Snyder, 1983). Trends in private and public enrollment should be monitored for white and minority groups. In 1982, 61% of 3- and 4-year-olds enrolled in school were in private programs. More than 70% of white but only 33% of Black children enrolled in school were in private programs (Kahdy, 1985). The high cost of good private day care may mean that poor children are denied access to quality educational experiences. These programs may be the beginning of segregated education for Black and white children, and for poor and affluent children. The provision of public programs for only poor children, however, will further institutionalize the segregation and unequal experiences of children of this age group. The adequacy of federal funds for poor children should be monitored. Federal Title XX and Head Start provide programs for only 20% of eligible children (Kahdy, 1985).

The reporting of data on 3- and 4-year-old enrollment should be in the most useful form. For example, 3- and 4-year-old enrollment combined with 5-year-old enrollment is misleading because almost 100% of 5-year-olds are in some school program. Another problem is that no distinction is made between half-day and full-day programs in the reports.

The cost of programs and the certification of teachers for this age group should be assessed. Day care workers generally earn less than public school teachers, have less formal education, and are not certified. On the other hand, public prekindergarten programs may be cost-effective if they employ unused or underused facilities and resources, and if existing services in the public schools, such as those related to counseling, testing, and health, can accommodate prekindergarten children. The instructional quality of programs should be assessed, especially the degree of structure of the curriculum, the nature of evaluation of students, and the attention given to social, emotional, and physical as well as cognitive development. The extent of before- and after-school programs in public and private institutions should be assessed, including the quality of children's experiences in those settings. The attitudes of the public, including private day-care providers, toward public prekindergarten programs should be assessed.

Adolescent development. Longitudinal studies of high school students should be continued. As adolescents become increasingly independent and begin to make decisions for themselves that may affect their school performance and their adult life chances, attention must be given to the special problems that arise. One is the relation of athletics to academic performance. Much attention has been given to young men, in particular, who are superb athletes in high school and later in college but whose academic skills are woefully lacking. School districts try to handle this situation by maintaining minimum academic standards for participation in athletic activities and other extracurricular activities. The athletics/academics problem appears to affect minority students more than whites. Career guidance in general is an important issue for Black males (Perry & Locke, 1985).

An additional problem of teenage students is alcohol and drug abuse. For example, a northern New Jersey school district recently established a policy requiring students to take a physical examination yearly, which includes tests for illegal drugs. Such extreme measures are controversial but reflect the concern with schools as places for drug use and drug trafficking. A related issue is violence in schools. Schools must have an atmosphere of safety rather than fear in order for teaching and learning to proceed smoothly. Cooke et al. (1985) point out problems in accurately assessing the incidence of student victimization.

The high incidence of pregnancy among adolescents is an issue for schools. Some schools provide sex education programs aimed at pregnancy prevention but, again, these are controversial. Many parents disapprove of the school's providing information about an aspect of children's lives that involves personal values. A 1982 survey found that 80% of approximately 200 urban school districts provided sex education but only 16% of senior high and 11% of junior high schools offered separate courses (Sonenstein & Pittman, 1984). Fewer than 10% of students attend formal programs of more than 40 hours (Kirby, 1984). Teen pregnancy has implications for the drop-out rate and may require school districts to provide special programs or schools for the pregnant teen. Among white females who dropped out of the sophomore class of 1980, the most frequently cited reason for leaving school (36.4%) was marriage or plans for marriage, which could have involved a pregnancy. Pregnancy was cited as a reason by 20.5% of white girls. For minority girls, the most frequently cited reason was poor grades (30%) but pregnancy was a close second (29.2%). Marriage or plans for marriage was cited by 19.2%.

Quality of instruction

Teacher preparation. The competence of teachers must be monitored in a more precise manner than merely assessing years of education or experience, although certainly those pieces of data should be collected. Much concern has been expressed regarding teachers' academic ability and the quality of courses they choose in college. In 1982, high school seniors planning to become teachers earned a low mean SAT score, ranking 26 out of 29 planned majors. College graduates who enter and plan to remain in the teaching profession are among the lowest scorers on the SAT (NCES Indicators, 1985). The Southern Regional Education Board's examination of transcripts of 1982-83 graduates of seventeen Southern universities revealed teachers take 65% more education methods courses than needed for certification, only 22% of their math courses are college-level, 75% take no foreign language and no philosophy courses, 66% take no chemistry and no physics, and 83% take no upper-level English courses (Galambos, 1985). Proposals to change the structure of teacher education have been made, such as having a more comprehensive undergraduate education and professional teacher training at the master's and doctoral levels. The specific structure of teachers' education needs to be monitored.

Teacher pay must be monitored. Teacher salaries, adjusted for inflation, have declined from the 1970's until the present (Weinberg, Gerald, & Tron, 1984). Male teachers' income is considerably less than the mean for all salaried professionals, although the comparable difference for female teachers is not nearly as great. In the two decades between 1961 and 1981, the adjusted mean salary for all full-time male workers rose 19% and for females rose 17%. In contrast, male teachers' adjusted salary rose only 2% and female teachers' slightly more than 1%. The prospect of low pay may discourage many talented college students from pursuing teaching as a career. Some efforts to remedy this situation have focused on strategies such as merit pay rather than across-the-board increases.

Not a new remedy, merit pay for teachers was tried in the 1920s and again in the 1960s; the programs instituted were short-lived (Johnson, 1984). In the present context of overall low teacher salaries, some educators view merit pay as impractical, divisive, and potentially unfair (Barranco, 1984). Although proponents of merit pay argue that all occupations reward extraordinary performance, merit pay actually is not used extensively in industry and is most effective in occupations, such as sales, that rely on individual effort and have clearly identified standards for success (Johnson, 1984). In contrast, effective teaching is poorly defined and depends on factors the teacher cannot control, such as the skills of administrators and of students' previous teachers (Johnson, 1984). Duckworth (1984) points out that the relation of beginning to maximum salaries for teachers is not much different from that of other salaried professionals, such as engineers; the striking difference is the overall low level of teachers' pay. A Boston public school administrator (Rosen, 1984), however, concluded that merely raising teacher salaries will not result in improved education, because Boston teachers have high salaries, relative to teachers nationwide, and their students very low achievement scores. Perhaps salary increases must be combined with other improvements in working conditions to affect teacher performance.

The supply of qualified teachers needs to be monitored and the racial/ethnic composition of the teacher pool needs to be monitored. In addition, demand for teachers needs to be accurately monitored so that efforts can be made to match supply with demand. The number of positions vacant can be underestimated if a program is terminated because of no teachers; no vacant or unfilled positions would appear because the entire program was terminated.

Classroom processes. The amount of time during the school day actually spent in instruction--time on task--has been conceptualized as a major variable of importance in students' achievement but no national data on this hard-to-measure variable are available (Cooke et al., 1985). Another important time variable, the proportion of students who attend school daily may not be accurate because states may compute attendance in different ways, and may even count students who have excused absences as present (Cooke et al., 1985). A third time variable, course enrollment, also may be inaccurate, especially if reported by students themselves (Cooke et al., 1985). Other time variables that need to be assessed are length of school day and school year, as some districts have experimented with this method of increasing instructional time.

Some educators (e.g., Barranco, 1984) point out that a more important issue is the quality of instructional time. The quality of instruction is even more difficult to measure than the time variables. Sirotnik (1983) found that teaching practices are remarkably similar across many elementary and secondary classrooms and are not substantially different from those employed throughout this century. These practices consist of teachers lecturing or students working on written assignments for the majority of class time. Questions typically were closed and factual, with little feedback or guidance. According to Sirotnik, the teaching practices supported dependence on authority, apathy, and passivity.

The nature of the interactions between teachers and students need to be monitored. Data suggest that teachers respond differently to minority and white children, to poor and middle-class children, to male and female children (see Lockheed, 1985). Differential interactions may affect the performance of children. The means by which teachers maintain order and discipline in classrooms should be monitored. A number of organizations, such as the American Medical Association, have issued formal statements opposing the use of corporal punishment in schools; four states have outlawed corporal punishment in schools. Suspending and expelling students from schools as disciplinary measures should also be monitored, especially among minority males, who are disproportionately subjected to these forms of punishment.

Public perceptions of education

Assessments of the views of parents and other citizens regarding education need to be continued and strengthened. Recently, Phi Delta Kappa established a National Commission on Public Confidence in Education. In 1970 and each year since then, a Gallup survey of public perceptions of schools, sponsored by Phi Delta Kappa, has been conducted. Although the 1984 results (Gallup, 1984; NCES Indicators, 1985) are hailed as demonstrating a sharp upturn in public opinion regarding schools, fewer than half of the 1984 respondents believed that the schools deserved a grade of A or B. The mean grade given schools in 1984 was slightly higher in 1984, but at 2.36 (with 2 representing C), compared to 2.12 in 1983, the claim of a sharp increase is hardly justified. Asking individuals what grade they would give their local schools may not be the most meaningful way to assess public opinion. The meanings attached to grades probably vary widely among individuals. Some may think that "C" is a good grade; others may believe it is a terrible grade.

The views of teachers and school administrators need to be assessed separately from those of parents. The 1984 survey of teachers conducted by Louis Harris Associates (NCES Indicators, 1985) found that teachers viewed student lack of interest, inadequate finances, and overcrowded classrooms as more serious problems than student discipline and drugs, the problems most frequently cited by parents in the 1984 Gallup Poll. A substantial number of teachers, however, viewed discipline (40%) and drugs (33%) as problems. The views of those with special needs or concerns vis-a-vis the school might also be assessed separately. Parents of handicapped children, gifted children, or learning disabled children may have different views of schools, as may minority and poor parents.

Summary.

National statistics should reflect the educational goals of the society, which are many and varied. The clear identification of major educational goals must be followed by careful operationalization of variables, uniform, commonly understood definitions, appropriate sampling techniques, and accurate reporting with checks for accuracy. Variables studied should be those that are most meaningful rather than merely those that are most easily quantified. More information is needed on the processes of schooling--what children actually experience--as well as on the outcomes of schooling. Research efforts of individuals and agencies should be coordinated as much as is possible. Specific needs include:

1. Measures of educational achievement that reflect the goals of the educational system.

2. Careful assessment and reporting of the educational experiences and achievement of underserved groups, including minority, poor, female, and handicapped students. Periodic special reports and oversampling of these groups may be required.

3. Assessment of bilingual programs, including appropriateness of criteria for inclusion; impact on achievement, retention of native language, and segregation; and appropriateness of teachers and curricula.

4. Assessment of private schools, including minority access, achievement, quality of curricula, teacher certification, and public support.

5. Assessment of computer use, including impact on learning and thinking, on curriculum and classroom processes, and on existing achievement gaps between various demographic groups.

6. Accurate assessment of the drop-out problem, including rates for various demographic groups, students' reasons for leaving school, and students' activities after leaving school.

7. Assessment of students' work experiences after high school, with attention to the nature and requirements of projected jobs.

8. Longitudinal studies of both elementary and secondary students, including achievement but also focusing on children as respondents regarding their own experiences and perceptions, and including problems in adolescence, such as pregnancy and drug abuse.

9. Careful assessment of public and private programs for 3- and 4-year-olds, including cost, curricula, teacher qualifications, length of school day, achievement, and enrollment of minority and poor children.

10. Assessment of teacher preparation, teacher pay, and teacher demand in various subject areas.

11. Assessment of classroom processes, including instructional time variables, teacher-student and student-student interaction, and measures of quality of instruction.

12. Meaningful assessments of perceptions of parents, teachers, and other citizens of the educational system, including the perceptions of parents of children with special needs.

Many of the problems of schools go beyond individual achievement, reflecting the developmental needs of children and the economic, political, and cultural conditions of American society. It is not sufficient to ignore issues larger than cognitive skills, using the defense that schools are not responsible. Whether the influence is intentional or not, schools affect children's overall development. And, intentional or not, the school is an arena in which the race/ethnicity, class, and gender biases extant in the larger society are played out. A critical issue for American education is whether schools can help all children develop into adults who have productive, personally satisfying roles in American society or whether schools will sort poor, minority, and female children into the roles traditionally allowed for them.

References

Alvarado, D. (1985, May 8). Home education ruled legal. *The Raleigh News and Observer*, pp. 1, 16.

Barranco, J. (1984). Accountable after the politicians have moved on. *Harvard Educational Review*, *54*, 6-10.

Bell, T.H. (1982). The federal role in education. *Harvard Educational Review*, *52*, 375-380.

Belsky, J., Steinberg, L.D. (1978). The effects of day care: A critical review. *Child Development*, *49*, 929-949.

Braddock, J. H. (1981a). The issue is still equality of educational opportunity. *Harvard Educational Review*, *51*, 490-496.

Braddock, J. H. (1981b). Quality and equality: Compatible or incompatible goals? *Phi Delta Kappan*, 166.

Caporael, L. R., & Thorngate, W. (1984). Introduction: Towards the social psychology of computing. *Journal of Social Issues*, *40*, 1-13.

Chaille, C., & Littman, B. (1985). Computers in early education: The child as theory builder. In E. L. Klein (Ed.), *Children and computers*. New Directions for Child Development, no. 28. San Francisco: Jossey-Bass.

Coleman, J., Hoffer, T., & Kilgore, S. (1981). *Public and private schools*. Chicago: National Opinion Research Center.

Culliton, B. J. (1985). Bok puts computers in their place. *Science*, *228*, 697.

Dickson, W. P. (1985). Thought-provoking software: Juxtaposing symbol systems. *Educational Researcher*, *14*, 30-38.

Duckworth, E. (1984). What teachers know: The best knowledge base. *Harvard Educational Review*. *54*, 15-19.

Duran, R.P. (1984). An information-processing approach to the study of Hispanic bilinguals' cognition. In J.L. Martinez, Jr., & R.H. Mendoza (Eds.), *Chicano psychology* second edition. New York: Academic Press.

Galambos, E. C. (1985). *Teacher preparation: The anatomy of a college degree*. Atlanta: Southern Regional Educational Board.

Gallup, G. (1984). The sixteenth annual poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, *66*, 23-38.

Garcia, E. E., & Gonzalez, G. (1984). The interrelationship of Spanish and Spanish-English language acquisition in the Hispanic child. In J. L. Martinez, Jr., & R.H. Mendoza (Eds.), *Chicano psychology* second edition. New York: Academic Press.

Giroux, H. A. (1984). Public philosophy and the crisis in education. *Harvard Educational Review*, *54*, 186-194.

Hallinan, M. (1982). Classroom racial composition and children's friendships. *Social Forces*, *61*, 56-72.

Hawkins, J., Sheingold, K., Gearhart, M., & Berger, C. (1982).

Microcomputers in schools: Impact on the social life of elementary classrooms. *Journal of Applied Developmental Psychology*, *3*, 361-373.

Harvey, W. (1985). Excellence and cultural pluralism: The challenge for comtemporary public education. *Teacher Education and Practice*, 47-50.

Howe, H. (1984). Introduction, Symposium on the year of the reports. *Harvard Educational Review*, *54*, 1-5.

Johnson, S. M. (1984). Merit pay for teachers: A poor prescription for reform. *Harvard Educational Review*, *54*, 175-185.

Kahdy, G. A. (May 1985). Prekindergarten programs in the public schools of North Carolina. Report to the State Superintendent of Schools.

Kirby, D. (1984). *Sexuality education: An evaluation of programs and their effects*. Arlington, VA: Mathtech, Inc.

Klein, E. L. (1985). Computer graphics, visual imagery, and spatial thought. In E. L. Klein (Ed.), *Children and computers*. New Directions for Child Development, no. 28. San Francisco: Jossey-Bass.

Klein, S.S.(Ed.) (1985). Handbook for achieving sex equity through education. Baltimore: Johns Hopkins University Press.

La Fromboise, T.D., & Plake, B.S. (1983). Toward meeting the research needs of American Indians. *Harvard Educational Review*, *53*, 45-51.

Lazar, I., & Darlington, R. (1982). Lasting effects of early education. *Monographs of the Society for Research in Child Development*, *47* (2-3, Serial No. 195).

Lepper, M. R. (1985). Microcomputers in education: Motivational and social issues. American Psychologist, 40, 1-18.

Linn, M. C. (1985). The cognitive consequences of programming instruction in classrooms. Educational Researcher, 14, 14-29.

Lockheed, M. (1985). Sex equity in classroom organization and climate. In S.S. Klein (Ed.), Handbook for achieving sex equity in education. Baltimore: Johns Hopkins University Press.

National Center for Education Statistics (January, 1985). Indicators of education status and trends. Washington, DC: Department of Education.

National Commission on Excellence in Education (1983). A nation at risk: The imperative for educaional reform. Washington, DC: U. S. Department of Education.

North Carolina General Statutes, 1979, Article 39, Nonpublic Schools, Chapter 115C-547-562 (1981).

Ogbu, J.U. (1978). Minority education and caste. New York: Academic Press.

Olson, D. R. (1985). Computers as tools of the intellect. Educational Researcher, 14, 5-7.

Otheguy, R. (1982). Thinking about bilingual education: A critical appraisal. Harvard Educational Review, 52, 301-314.

Padilla, A. M., & Lindholm, K. J. (1984). Child bilingualism: The same old issues revisited. In J. L. Martinez, Jr., & R. H. Mendoza (Eds.), Chicano psychology second edition. New York: Academic Press.

Pea, R. D. (1985). Integrating human and computer intelligence. In E. L. Klein (Ed.), Children and computers. New Directions for Child Development, no. 28. San Francisco: Jossey-Bass.

Perkins, D. (1985, May 9). Court ruling to lead to more `home schools', officials say. The Raleigh News and Observer, pp. 1, 4.

Perry, J.L., & Locke, D.C. (1985). Career development of black men: Implications for school guidance services. Journal of Multicultural Counseling and Development, 13, 106-111.

Reid, P.T. (1982). Socialization of black female children. In P.W. Berman & E.R. Ramey (Eds.), Women: A developmental perspective. Washington, DC: U.S. Department of Health and Human Services.

Rosen, R.V. (1984). For the short and long term. Harvard Educational Review, 54, 23-27.

Rotberg, I. C. (1982). Some legal and research considerations in establishing federal policy in bilingual education. Harvard Educational Review, 52, 149-168.

Scott-Jones, D., & Nelson-Le Gall, S.N. (in press). Defining Black families: Past and present. In E. Seidman & J. Rappaport (Eds.), Redefining social problems. New York: Plenum.

Sheingold, K., Hawkins, J., & Char, C. (1984). "I'm the thinkest, you're the typist": The interaction of technology and the social life of classrooms. Journal of Social Issues, 40, 49-61.

Sheingold, K., Kane, J. H., & Endreweit, M. E. (1983). Microcomputer use in schools: Developing a research agenda. Harvard Educational Review, 53, 412-432.

Sinclair, E. (1983). Bilingual education: Fact or fancy? In G.J. Powell (Ed.), The psychosocial development of minority group children. New York: Brunner/Mazel.

Sirotnik, K.A. (1983). What you see is what you get: Consistency, persistency, and mediocrity in classrooms. Harvard Educational Review, 53, 16-31.

Slaughter, D. T., & Schneider, B. L. (1985, April). Assessing the educational goals of private school parents: Comparing Black and nonblack families. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Slaughter, D. T., Schneider, B. L., & Lindsey, R. W. (1985, April). The relation of family-school goals to the educational experiences of Black children in private schools. Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Snyder, T. D., & Grant, W. V. (1983). Digest of education statistics 1983-84. Washington, DC: U. S. Government Printing Office.

Sonenstein, F. L., & Pittman, K. J. (1984). The availability of sex education in large city school districts. Family Planning Perspectives.

Stage, E.K., Kreinberg, N., & Eccles, J. (1985). Increasing the participation of and achievement of girls and women in mathematics, science, and engineering. In S.S. Klein (Ed.), Handbook for achieving sex equity through education. Baltimore: Johns Hopkins University Press.

Steinberg, L., Blinde, P.L., & Chan, K.S. (1984). Dropping out among language minority youth. *Review of Educational Research*, *54*, 113-132.

U.S. Department of Education (1985). State Education Statistics.

Weinberg, A., Gerald, D., & Tron, E. (1984). Elementary/Secondary education. In V. W. Plisko (Ed.), *The condition of education, 1984 edition*. Washington, DC: U.S. Government Printing Office.

Whalen, R. E. (1984). Secondary education: Student flows, course participation, and state requirements. In V. W. Plisko (Ed.), *The condition of education 1984 edition*. Washington, DC: U. S. Government Printing Office.

Wolf, D. P. (1985). Flexible texts: Computer editing in the study of writing. In E. L. Klein (Ed.), *Children and computers*. New Directions for Child Development, no. 28. San Francisco: Jossey-Bass.

Zill, N., Peterson, J. L., & Moore, K. (September, 1984). Improving National Statistics on Children, Youth, and Families. A report on recommendations made at the Interagency Conference on Child and Family Statistics. Washington, DC: Child Trends.

June, 1985

"Educational Indicators: What Do We Need to Know That We Don't Know Now?"

Prepared for the National Center for
Education Statistics by

Ramsay W. Selden,
Excellence in Education Program,
National Institute of Education

When the National Commission on Excellence in Education prepared its report, A Nation At Risk, it was concerned very much with educational indicators. The Commission was charged, in part, with determining the nature of the major problems facing American education. To do this, the Commission needed to examine the evidence that was available of the health of the country's educational system. The products of this effort are two sections in particular in A Nation At Risk: a section called "Indicators of the Risk," which is a compendium of the major pieces of evidence the Commission found concerning the quality of American education; and sections of "Findings" in the areas under which the Commission organized its analysis and recommendations--time, content, expectations, teaching, and leadership.

The figures cited in these sections encompass student achievement in basic skills and academic subjects, rates of functional illiteracy, trends in the amount of homework assigned to students, figures on average teacher salaries, comparisons of the time spent in school by U.S. students with the time spent by their counterparts in other countries, and so on, but the Commission's experience revealed perhaps as much about the inadequacy of our educational indicators as it did about the inadequacy of our educational programs. The "Findings" and "Indicators of the Risk" cited by the Commission were the best evidence available on important aspects of our educational system, but there was much more information that should have been included in these sections that simply was not available. The Commission could say nothing, for example, about the general skill of American teachers in presenting subject matter and conducting lessons. No figures were available to indicate whether functional illiteracy was increasing or decreasing over time. Comparisons of the achievement of U.S. students with that of students in other countries was fifteen years old for some subject areas. Finally, detailed information about what high school students actually studied and knew in subjects like science, social studies, mathematics, literature, and the humanities had never been collected or reported on a nationally-representative sample.

At least one indicator--the courses typically taken by high school students during their four years in high school--was felt to be so important that the Commission arranged for its collection, having found it was not available from any other source. The resulting study revealed that students took alarmingly few courses in the academic subjects. This finding led the Commission to recommend that a minimum number of courses be required in the academic "basics"--English, science, social studies, and mathematics, as well as computer science and (for the college bound) foreign languages--in order for students to be granted high school diplomas, and over the past 3 or 4 years, most states have increased these course requirements. This experience underscores the importance of this information, a level of importance which is not consistent with the inattention that this particular piece of information has been given among our educational data-collection activities.

Significant efforts are underway to improve our educational indicators. The Department of Education initiated an "indicators" project that has, thus far, provided a useful taxonomy for educational indicators and a description or compendium of twenty of the key indicators in this taxonomy that are currently available. A periodic report of these key indicators was initiated by NCES in 1985. (Indicators, 1985) The Department also initiated a one-sheet chart displaying key educational information for each of the states. (U.S. Department of Education, 1984, 1985) The Chief State School Officers as a group have endorsed (and taken the initiative for compiling) educational assessment data, including school outcome data, on a state-by-state basis, and the National Assessment of Educational Progress has taken steps to enable states to piggy-back on the National Assessment to collect state-by-state achievement data for purposes of self-monitoring and comparison.

Despite these efforts, a crucial step remains to be taken. This step is to consider systematically the information that we need but do not have about education, identifying and planning those indicators that should be added to, or improved among, our current set.

What I will attempt to do in the remainder of this paper is to establish a framework for organizing and considering educational indicators, review which of the indicators in this framework are currently available, and identify some of the important indicators that, based on the experience of the Commission on Excellence, are not available.

A Framework for Educational Indicators

The concept of an educational indicator involves two elements. Each of these has been addressed by other analysts, but they have not generally been considered together. First, an indicator describes a variable in the educational system or educational process. The National Commission on Excellence in Education organized these variables under an input-output model, with an emphasis on the inputs of education:

Figure 1

Framework for Variables in Education Used By
The National Commission on Excellence in Education*

INPUTS					OUTPUTS	
Time	Content	Standards & Expectations	Teaching	Leadership & Support	Achievement	Retention
Days/yr; Hrs/day; Use of time	Courses taken; Curric. content	Grades; Test scores for promotion; Rigor of materials; Stds for admission; Homework assigned	Salaries; Size of workforce; Aptitude scores; Rewards for performance	Fiscal support; Community initiatives	Test scores; NAEP results Int'l comparisons	Drop-out rates; Levels of attainment

*The variables shown are intended to illustrate the categories in this framework; they are not necessarily all of the variables considered by NCEE.

The Education Department's Indicators Project uses a similar scheme, describing education as a process, but it classifies the indicators under "outcomes," "resources," and the "context of education," or those factors such as instructional climate and parent support that have an effect on the success of the educational process.

These schemes are alternatives that attempt to serve the same purpose: to model the educational process or system. One must decide if the models are interchangeable or whether they differ, and, if they differ, whether one of them is more valid, comprehensive, or parsimonious than the others.

The second element involved in the idea of an educational indicator is the purpose to which the indicator is applied (Selden, 1984; and Smith, 1984). As we have pointed out, indicators can be compared to absolute standards of how we want the educational system to perform. They can be used to compare the performance of our system with the performance of other systems. Or, they can be used to see how our system is doing in relation to how it has performed in the past. The Commission looked at evidence of the quality of American education by comparing it in some cases to absolute standards (23 million illiterate adults is more than we should accept in this society); by comparing it with similar evidence pertaining to education in other countries (twelfth graders in the U.S. do fair to poorly in mathematics when ranked among twelfth graders from developed countries); or by comparing it with evidence of how we have done in the past (17-year-olds are less able to draw inferences from their reading now than they were in 1970). These are the three major types of analysis that can be made with an educational indicator.

Smith explains these functions quite ably and how they turn simple variables about education into indicators which are useful for setting policy. He also points out that indicators can be analyzed in conjunction with one another to explore how the educational system works (and how policy decisions might affect it). For example, while the aptitude of students entering teacher-preparation programs has declined, so have average teacher salaries, in real dollars, suggesting that there may be a relationship between these trends.

Given these two notions essential to the concept of an educational indicator (the features of the educational system that they describe, and the purposes for which they are examined), one is tempted to array indicators in a matrix:

Figure 2

Matrix of Educational Variables and
The Purposes to Which They Can be Put As Indicators

	EDUCATIONAL SYSTEM	
PURPOSES	Inputs	Outputs
Comparison with a standard	XXX	XXX
Comparison with another system	XXX	XXX
Comparison with past performance	XXX	XXX
Analysis of how the system works	XXX	XXX

In this matrix, each characteristic of the educational system could be reviewed for any or all of the four purposes shown, and any type of comparison or analysis could be made on any of the characteristics of the system. In real practice, it would not be useful or appropriate to make all of these analyses. Some of the inputs and outputs of the educational system have no inherent, absolute standards associated with them, or they may have no standard to which we would all agree. In these cases, reference to other systems or to past performance may be the only reasonable basis for interpreting the indicator. Examples of this situation might be indicators of the ability of teachers to present lessons before a classroom, or the fiscal effort communities put forth to pay for their schools. In these cases, we really have no basis for deciding the level on the indicator that is desireable or acceptable. In other cases, we may have set absolute goals for the U.S. educational system that are one-time efforts, or that are unique to the U.S. system and do not pertain to other systems, such as extending education to the disadvantaged over a certain period of time. In these cases, we would concentrate on our attainment of the goal and would not (perhaps could not) review progress in relation to past performance or to other countries.

With these caveats in mind, one can imagine a list of variables, and purposes to which they can be interpreted, that would constitute a comprehensive and "idealized" list of educational indicators. The question before us is: Among the most important indicators in such a list, which ones are available to us and which ones are not?

Educational Indicators--What Do We Know and Not Know?

The experience of the National Commission on Excellence in Education gives us a basis for identifying a preliminary set of some of the indicators that would be useful to policy makers, and that are not now available. To do this, let us go through the taxonomy of indicators that the Commission considered, to identify at least some of the analyses that I feel the Commission was not able to make, because the necessary data did not seem to be available or useful.

INPUTS

1. Time--The amount of time allocted to education, and the use of the time allocated by teachers and students.

What the Commission reported concerning time:

° That, in the U.S., the typical school year is 180 days and the typical high school day is 6 hours, compared with England and other industrialized countries, where it is not unusual for high school students to spend 8 hours per day, 220 days per year in school. (A Nation At Risk, page 21)

° That, out of 30 hours of student attendance in school each week, the average school provides 22 hours of academic instruction, and some schools provide only 17 hours. (Risk, page 22)

° That, because of differences among teachers in their ability to manage classroom time, some students may receive only one-fifth the instruction in reading comprehension that is provided to other students. (Risk, page 22)

° That, in most schools, students are not taught how to manage their time, through planned and systematic instruction in study skills. (Risk, page 22)

° And that students in the U.S. are asked to spend too little time on homework, and the time they spend on it has declined. (Risk, page 19)

What the Commission could not report...

...concerning the volume of time allocated to schooling:

° How much states and local school districts vary in the length of the school day or school year: for example, how many districts provide five, six, or seven hours of attendance per day, or what proportion of districts offer more or fewer than 180 days of instruction per year?

° Whether the length of the typical school day or school year has been increasing or decreasing over time in this country.

...concerning the use of time by teachers and students:

° Whether (and how much) schools around the country, at both the elementary and secondary levels, vary in the proportion of allocated time that they actually schedule for academic instruction.

° How U.S. schools compare with schools in other countries in the proportion of allocated time that they schedule for academic instruction.

° How U.S. teachers compare with teachers in other countries in their ability to manage classroom time.

° Whether teachers have become more or less proficient over the years in their ability to engage students in academic learning.

° Whether students in other countries are taught work habits and study skills better than students are taught these skills in this country.

° How much variation exists in the teaching of study skills among schools in this country, and whether the variation is systematically related to characteristics of the schools. (Do certain types of schools provide this kind of instruction?)

° And, how U.S. students compare with students in other countries in how much homework they are required to do.

2. Content--the courses taken by students, or the subject matter presented to students, and the substance learned in these courses.

What the Commission reported concerning content:

° That high school students have migrated to the general track from vocational and college-prep programs, resulting in a lack of focus in their curricula. (Risk, page 18)

° That few high school students were taking courses in such subjects as advanced math, foreign languages, or geography, even though these courses have been offered to them, generally.* (Risk, page 18)

° And that too many of the credits being earned by high school students have been in courses in health and physical education, out-of-school work experience, remediation, and courses aimed at personal service or development, such as training for adulthood and marriage.* (Risk, page 18-19)

(The Commission addressed the courses taken by high school students under two areas--under "Content," concerning the nature of the coursework students take or do not take, and under "Standards and Expectations," reflecting the rigor of the programs of coursework we expect of students. Indicators of the nature of the coursework taken by, or presented to, students are discussed here.)

What the Commission could not report concerning content:

° The nature of the courses taken by students in high school.*

° The content of the courses taken: for example, what students study in "General Science" or "World History" in high school, or in English in the fifth grade.

° Whether the set of courses typically taken by students in high school has changed over time, beyond shifts among tracks or programs.

° How the substance of the courses offered to, and taken by, students in high school, or the substance of subject matter taught in elementary school, has changed over time.

° Whether the courses taken by students in high school in this country differ from the courses taken by high school students in other countries.

° Whether, and how, the substance of the courses typically offered and taken by students in high school, or the substance of the subject matter taught in elementary school, differs among industrialized countries.

---

*The Commission was able to report on students' course-taking patterns only because it requested and supported a one-time survey of the transcripts of high school graduates. The information was not then readily available among data-reporting activities, and still is not.

° How the content of courses taken by students in high school, or of subject matter taught in elementary school, corresponds to what the public and its educational leaders feel the schools should teach.

° Whether the material in widely-used textbooks is up to date.

3. Standards and Expectations--The requirements and values communicated to students at points in their school careers.

What the Commission reported concerning standards and expectations:

° That, during a period of time when the amount of homework assigned to students was declining, grades rose and achievement fell. (Risk, pages 19-20)

° That students in other countries were required to take three times as many advanced mathematics and science courses as students in the U.S. (Risk, page 20)

° That states required too few courses in the academic basics of mathematics, science and computer technology, social studies, English, and foreign languages. (Risk, page 20)

° That statewide testing programs aimed almost universally for "minimal" competencies, minimums which had become the "maximum" expectations that were communicated to students. (Risk, page 20)

° That many public colleges have changed their admission policies to accept any high school graduate in their state, serving notice that the nature of a student's coursework or his or her grade point average in high school make no difference in determining whether he or she attends college. (The Commission noted that a trend in this direction also had taken place among more selective colleges and universities.) (Risk, pages 20-21)

° And that the textbooks and other instructional materials used in school demand too little of students, having been written down to lower reading levels, showing less and less of the influence of experienced teachers and scholars, and presenting students with academic material that would not be likely to challenge many of them. (Risk, page 21)

What the Commission could not report concerning standards and expectations:

° Whether the courses required by states for graduation from high school have changed over time, and, if so, how.

° Whether local course requirements for high school students differ from state requirements, and, if so, what they are and how they differ.

° Whether local requirements are changing over time, and, if so, how.

° The rigor of the school work that is actually required of students through assignments, homework, and course tests and quizzes, beyond the inferences that can be made from textbooks and achievement tests.

° Whether and how the academic demands of school work have changed over time, and how the work done by students in this country compares with the work required of students in other countries.

° Whether, in general, the content of textbooks has become less demanding, academically, over time, or whether the books have just become easier to read.

° Whether state and local testing programs are becoming more or less demanding, over time.

° What standards or expectations grades or teacher testing in subject matter communicate to students, and how these teacher standards or expectations compare with local or state achievement testing and minimal competency testing.

° How the standards represented by grades and teacher testing in this country compare to grading and subject matter testing of teachers in other countries.

° Whether grading has become "inflated" over time--that is, whether teachers are giving higher grades for comparable work, compared with grades given in the past.

° Whether there has been a trend up or down over time in the use of tests or other standards to determine if a student is promoted from grade to grade or level to level in school.

° Whether the school systems of other countries use tests and other yardsticks more than we do to determine progress through the school system.

° Whether parents expect more or less from their children in school now than they used to expect and communicate.

4. Teaching--the quality of teaching as a practice and as a profession.

What the Commission reported concerning the quality of teaching...

...as a practice:

° That the aptitude test scores of students going into teaching are too low. (Risk, page 22)

° That too little of the teacher preparation program, especially for the preparation of elementary school teachers, consists of courses in educational methods, and too little in the subjects to be taught. (Risk, page 22)

...as a profession:

- That teachers are paid too little and have too little responsibility for important professional decisions, such as the selection of the textbooks they use. (Risk, pages 22-23)

- That there are critical shortages of teachers in certain subjects (mathematics, science, and foreign languages) and specialties (education of the gifted and talented, of language minority students, and of the disadvantaged). (Risk, page 23)

- And that shortages of teachers in mathematics, science, and English are resulting in a large proportion of new teachers hired to teach in these areas who are not trained or certified to teach the subjects. (Risk, page 23)

What the Commission could not report concerning the quality of teaching...

...as a practice:

- The capability of our teachers, directly measured, in the professional practice of planning and conducting lessons.

- The pedagogical skill, generally, of our teachers, compared with the teachers of the past, or with the classroom skill of teachers in other countries.

- The quality of teachers' knowledge of the subject matter they teach, measured against notions of what they should know, compared with teachers in the past, or compared with teachers in other countries.

- The ability of our teachers to teach academic subjects to students who vary in background, capability, and interest or motivation; whether our teachers are improving in this ability over time; and how our teachers compare in this regard with teachers trained in other countries.

- The relationship between alternative approaches to teacher preparation and the relative proficiency of teachers in the classroom.

- The prevalence or status of different approaches to teacher preparation that seem to be either successful or unsuccessful.

- The ability of teachers to handle classroom discipline problems; how this ability has been changing over time; and how our teachers compare in this regard with teachers in other countries.

- The nature and success of efforts by school systems and states to improve the proficiency of teachers through staff development or inservice training.

What the Commission could not report concerning the quality of teaching...

...as a profession:

° The range in average teacher salaries among states, school districts, and professional specializations.

° How the salaries of U.S. teachers compare with those of teachers in other countries, in terms of relative buying power, and whether the buying power of teachers is going up or down in different countries.

° Early indicators of the long-range supply of teachers, such as the career intentions of students graduating from high school.

° Reasons cited by high school seniors or college students for not going into teaching (ie, low salaries, the poor prestige of the profession, perceptions of teachers having little professional autonomy, poor working conditions, etc.).

° The relative importance of various incentives in making teaching more attractive to those in the profession, such as higher salaries, greater possibility for career growth, more professional autonomy, better working conditions and support.

° The status of various professional incentives for teachers, both in this country over time and in other countries.

° Systematic tracking of the demand for, and supply of, teachers broken down into various professional specialties, and reported for states or regions of the country.

° Trends over time in the standards applied by states to certify new teachers and to maintain the certificates of veteran teachers.

° The nature of the standards applied by local school districts in recruiting and retaining teachers, including the use of different approaches for evaluating teacher performance.

° How the public perceives the status of teaching as a profession, and how this has changed over time.

° Perceptions of the status of teaching as a profession in this country, compared to its prestige in countries with different histories, cultural values, and organizational structures for education.

5. Leadership and Support--Efforts by the public, policy-makers, and parents to support and provide direction to the schools.

(The Commission on Excellence did not report findings having to do with the support and leadership provided to the schools, but the Commission did make recommendations in this area, implying the need for the following indicators, which do not seem to be available, now:)

° Regular reports of the fiscal effort put forth by different countries to support their schools, in terms of school expenditures as a proportion of GNP.

° Evidence of the range and level of specific efforts to support schools at the state and local levels, such as levels of expenditure, number or proportion of bond issues passed, or changes in state appropriations for education.

° The range in tax burden for education among states and local school districts.

° The relative cost of meeting different educational demands, such as education of disadvantaged, handicapped, or gifted and talented students.

° Geographical differences in the costs of education, such as the costs among regions of the country or different costs for schooling in sparsely-populated areas, cities, and other types of community.

° The relationships between cost factors and educational outputs.

° The involvement of parents and the local community in making decisions about school programs, or in contributing to these programs.

° The nature and level of efforts by parents to monitor and encourage the progress of their children in school.

° The nature and number of cooperative programs beween business and the schools.

° The perceptions of policy-makers and leaders at the national, state, and local levels about the quality of the schools and what the strengths and weaknesses of the schools may be.

° The nature of school improvement efforts being launched at the national, state, and local levels, and trends in these efforts over time.

OUTPUTS

What the Commission reported concerning student achievement (Risk, pages 8-9):

° That there had been general patterns of decline in scores on standardized achievement tests administered nationwide.

° That scores had declined in general on the Scholastic Aptitude Tests.

° That scores had declined on the College Board achievement tests in subjects such as physics and English.

° That achievement in science had declined steadily on the National

Assessment of Educational Progress from 1969 to 1977.

° That there were specific problems of poor (and declining) performance in the "higher order" aspects of reading, writing, and mathematics on the National Assessment.

° That U.S. students fared poorly in international comparison of achievement.

° That the number and proportion of students scoring 650 or higher on the SAT's had dropped.

° That most gifted students appear to be achieving at a level below their tested aptitude.

° That scores on the Graduate Record Examination had declined.

° Results of one-time surveys of functional illiteracy which indicated that about 23 million adults were functionally illiterate, that about 13% of school-aged youth were functionally illiterate, and that over 40% of minority youth were functionally illiterate.

The Commission also recognized (Risk, page 11) that the average person in the U.S. today is provided more schooling (if not better schooling) than the average person received a generation ago. This point is based on steady increases in the proportions of people completing high school, attending college, or completing college over the past 80-100 years.

What the Commission could not report concerning student achievement:

° The number or proportion of students nationwide who seem to know concepts and principles that would be recognized as basic to the academic subjects taught at different levels in school. For example, how many students understand and can apply the experimental method in science, know the central theme of Moby Dick, or can apply tenets of the Bill of Rights to contemporary situations?

° The achievement of today's students in the U.S. compared with contemporary students in other countries, as opposed to comparisons made 10-15 years ago.

° The ability of students in the U.S. to solve complex problems in the different academic subjects by finding and interpreting the appropriate information, reasoning analytically, and expressing their conclusions effectively.

° Trends over time on a comprehensive index of our educational productivity that would be more meaningful than SAT scores; for example, Wurtz and I (Wurtz and Selden, 1985) have suggested an annual "national educational index" based on the product of two numbers: the proportion of a common core of academic content objectives that students seem to have learned each year, and the percentage of students who have complete certain levels of schooling, such as high school, that year. (Such an index could, and should, also be adjusted for costs and educational effort factors over time and among school systems.)

° Trends over time in basic measures of educational achievement, such as functional illiteracy.

° The range in average student achievement among significant political units, such as states or a sample of local school districts.

° Trends in achievement over time in states or in local school districts facing different challenges and situations.

(Since the Commission did not consider many individual indicators of student participation and retention in school, and since these are areas where existing data are relatively extensive and useful, I will not devote space to consider what indicators of participation may not be available.)

Why We Do not Know Some of the Things We Need to Know About Our Schools

Given this preliminary, crudely-organized list of some of the indicators I believe the Commission could have used, but that I do not believe were available to it, let me conclude by considering why these indicators may not be available, so we can at least point toward developing them in the future.

First, let me explain the tests that I did and did not apply to this list. I have tried to list only indicators that are relatively significant: that is, they involve variables and analyses that would be particularly useful to educational decision-makers and policy-makers. Second, I have relied on my knowledge of the data on the educational system that are and are not available on a regular basis. I have not been able to conduct exhaustive searches to verify my understanding in every case. In some cases (not many, I hope), I may not be informed well enough; in other cases we may agree that related data are available, but I will believe that they are not fine-grained enough, or regular, enough to be useful.

INPUTS

Time. The status of indicators concerning the amount of time scheduled and allocated to schooling reveal a problem that will recur frequently in this discussion. The length of the school year, the length of the school day, and the structure of the school day are set by states, more or less officially depending on the state. Local school districts and individual schools, public and private, may meet, fail to meet, or exceed these standards, so that local practice may vary substantially. Only a census of the states and a regular survey of local districts would reveal norms and the degree of variability around the norm for these dimensions, and each brings with it effort in terms of data collection and reporting.

Indicators involving how time is used in school introduce the second problem endemic to indicators. Even though we know that how teachers use scheduled time is important, measuring this variable validly and monitoring it over time or among school systems present serious technical challenges, and would probably be costly. Similarly, measuring and tracking how well students are taught to use their own time to enhance learning would be difficult and costly.

Content. The problem here seems to be the level of detail of the information we collect. We are just beginning to monitor the courses that are required by states and local school districts for graduation from high school. While these efforts have been one-shot projects conducted to monitor recent reform initiatives, they could be made regular quite easily. The bigger challenge is collecting information on what students study below the level of the course: what is typically presented to students in each subject at each level of school, and how much does this content vary among schools around the country? This would require surveys of state and local curriculum objectives, content analyses of widely-used textbooks, and surveys of teachers to describe what other content and activities they present to their students.

Standards and Expectations. Needed here are surveys of state, local, and classroom practices in setting, communicating, and enforcing academic standards for students. We know how many states have minimal competency testing programs, and we recently learned what subjects they cover, at what levels they are administered, and how they are used, in most states. We also know, from occasional surveys, how much homework students are assigned. We do not know what standards states and local school districts apply to promote students from level to level. We do not know what criteria teachers apply in assigning grades and how they might be changing over time. We do not know how teachers use classroom tests of subject matter. We do not know enough about what parents expect of their children in school, or how well these expectations are communicated to students. In order to obtain this information, we would need new or expanded surveys of accreditation bodies, local school administrators, teachers, parents, and students.

Teaching. Indicators of the quality of teaching as a practice suffer one, central shortcoming: they are all indirect. We infer the competence or professional skill of teachers from aptitude test scores, college grades, courses studied, paper-and-pencil qualifying examinations, and compliance with certification standards, but not from direct measures of the ability of teachers to teach students. To measure teaching skill directly, we would need an informed definition of the qualities and behaviors that go into good teaching, and then we would need to operationalize this definition (or multiple definitions) with procedures for observing teachers. With definitions and operating procedures (which some states and many local school districts are developing in order to evaluate teachers) a national, or international, sample of teachers could be observed periodically to provide longitudinal and comparative data on the overall pedagogical ability of teachers. With these techniques, we could also refine the observational methods to measure special abilities of teachers--the ability to teach students with particular characteristics with which we are concerned, or to do other, specific aspects of the job of teaching.

Indicators on teaching as a profession will probably fill in quickly. There is a great demand for information on what state and local school systems are doing to enhance incentives and working conditions for teachers, covering salaries, pay-for-performance provisions, career ladders, professional evaluation systems, and steps to reduce non-professional duties and stress. The challenges for indicators in this area seem to be arriving at standard definitions of concepts such as salaries, benefits, and professional incentives, and then setting up routine programs and procedures for collecting and reporting this information on states and local school districts.

Leadership and Support. The Commission's analysis, and the school improvement efforts that came with the reform movement of which the Commission was a part, included many recommendations for local school leaders and administrators to enhance the involvement and support of parents, citizens, businesses and others in the schools. To track these efforts and at least their perceived effectiveness, surveys would be needed of school administrators, school board members, business leaders, and others to determine the extent and nature of efforts they have made to develop community support and to exercise new levels of leadership for the schools.

OUTPUTS (Achievement)

Currently-available achievement indicators are lacking in their level of detail. The National Assessment of Educational Progress has made a major contribution in providing us with periodic, national data at three age levels on student achievement in each of the basic academic subjects. What we need beyond this is to break down achievement data into greater detail. We need to monitor the specific portions of the core, academic subject matter we are concerned with in the schools that students know and do not know. NAEP is being augmented to assess the status of this kind of specific, subject-matter knowledge in the areas of literature and American history. Collecting this level of information across subjects on a regular basis would permit more effective fine-tuning of educational programs at the state, local, and classroom levels, allowing us to attend to those parts of the academic program that students do not seem to be learning. We cannot do this with the level of detail offered by NAEP's current design. We also need more detailed sampling and reporting, so that outcomes on the Assessment can be reported by state and by meaningful types of local school district. Finally, we need to collect information on educational programs or efforts that are associated with student achievement, in order to begin to understand what educational efforts or approaches seem to provide the best results; NAEP has begun to collect this kind of information on the schools it samples, and this effort should be refined and developed.

International indicators of achievement suffer mainly from infrequency. The IEA studies are on a 10- to 15-year cycle, meaning that sometimes the most recent available comparative data are very old. Five years would seem to be a reasonable time period for this cycle, to ensure that comparsions are not erroneously extrapolated from situations that no longer exist. Along with a shorter cycle, IEA-type studies should be made more routine. Under the present arrangements for conducting the IEA, we cannot count on the studies being repeated in the future; each cycle is conducted through private organizations in an ad hoc way. The program should be institutionalized so that participation in the studies and support for them can be relied-upon.

Summary

I have attempted to do three things in this paper. The first was to establish a structure by which one could talk about educational indicators--both those that we already know about and those that we may need, but do not have available to us, now. The second was to describe some of the indicators that at least one group, the National Commission on Excellence in Education, was unable to analyze because they did not, and still do not, seem to be available. The third was to speculate briefly on what would be involved in developing at least some of the important indicators that we are missing.

This paper should really be used a pilot-test for a process that should be done more thoroughly and systematically, but it does reveal both that there are important kinds of information about education that we lack, and that we can identify and address these gaps. Venturing forward into these undeveloped areas will take time, effort, and money, but few tasks offer as great a potential to give us useful tools for improving education. Information is power, and better information about education would give us tremendous leverage in managing and improving it.

Notes

A Nation At Risk: The Imperative For Educational Reform. Report of National Commission on Excellence in Education. Washington, D.C.: U.S. Department of Education, 1983.

Indicators. Washington, D.C.: National Center for Education Statistics, January, 1985.

Selden, Ramsay, "Lessons From National Assessments of Achievement," Presented to the Northeast Regional Meeting, Comparative and International Education Society, Washington, D.C., April, 1984.

Selden, Ramsay, and Emily Wurtz, "An Immodest Proposal For An Index Of Our National Educational Product," Staff paper, Excellence Indicators Division, National Institute of Education, February, 1985.

Smith, Marshall, "A Framework for the Development of National Educational Indicators," Prepared for the Council of Chief State School Officers, October, 1984.

U.S. Department of Education. "State Education Statistics," Department of Education, Office of Planning, Budget and Evaluation, 1984, 1985.

Improving the Quality and Usefullness of NCES data

Marshall S. Smith
University of Wisconsin, Madison

This paper sketches seven activities that NCES should pay attention to over the next few years as it attempts to provide accurate and complete data on the elementary and secondary education system of the United States. No attempt is made in this discussion to be comprehensive and no logical or programmatic priority should be inferred from the order of items on the list.

1). Act aggressively to meet the major problems of data collection and establish a system to monitor the quality of NCES data. Over the past year there have been a number of internal (Dept. of Education) and external (GAO, "The Sorry State...") critiques of various aspects of data gathered and reported by NCES. The time is right to dedicate resources to meet these problems. One approach is to deliberately review the critics' points and make adjustments to meet each of the specific criticisms. Another approach is to work with the NAS committee reviewing NCES to examine areas that need improvement and to suggest ways of implementing corrective action. Both approaches and others should be used and a clear document spelling out what NCES intends to do and is doing about the quality of their data should be developed and released to the field for comment. Where resources are not available or where policy (legislative or administrative) gets in the way of improving the data collection this should be pointed out in the document and a strategy indicated for meeting the problems.

Beyond correcting the immediate problems three other steps should be taken. First, I have a hunch that there is a great deal of unnecessary and useless data collected because of legislative or administrative mandate or because it has been collected in the past and no-one has gotten around to examining its continued utility. It would be useful to ask the Academy and internal NCES staff to suggest candidates for reduction in data collection. Second, you should establish a system of yearly review. I don't have in mind anything fancy. One approach would combine internal and external review. For example, internally to NCES, on a rotating basis each year a Director of one of the NCES programs could have an assignment of producing a planning document about ways of improving data collection and data use in the agency. Or, internally to the department, each year the planning office might be requested to review a different aspect of the data collection program and use of the data in the Department. On the outside you might contract with a single particularily knowledgeable person each year to produce a short (25 page) provocative paper on one or another aspect of the data collection and analysis activities of the agency. These papers might look both at present problems and future opportunities. The effort to gain outside advice (beyond the Advisory Committee) should not end with the effort of which this paper is a part.

Third, it should be possible to establish an ongoing system for partial verification of NCES data. One approach would be to use the Quick Response survey to cross-validate certain statistics each year. Another strategy would be to work with certain key states each year to develop cross-validating systems of data collection. I do not think that you should engage in massive cross-validation. The effort should be limited and selective. Major problems signalled by discrepancies which occurred in the cross-validation would require larger efforts to correct but the frequency of major problems should be substantially reduced by such an on-going attempt to insure accuracy.

2). Improve and coordinate the tests of academic achievement used in major surveys. One of the major problems with the HSB survey is the quality of the student outcome measures, particularily the student academic achievement tests. I am mindful of the robust psychometric properties of many of the tests -- as ETS has shown in their recent report. I also understand the constraints imposed by attempting to have direct comparablity among different surveys (eg. the NLS and the HSB and presumably the NLS, HSB and HSB2) and by the need to minimize the time spent by the students taking the tests. One component of the problem rests with the lack of relationship between the content of the tests and the curriculum of schools. One version of this criticism comes from content specialists eyeballing the tests and claiming that there is little relationship between the tests and the curriculum. For example, there is little attention paid to the content of English and History--the subjects in high school in which students spend a large percentage of their time. And the tests assess only a limited conception of math and science. Another related version of the criticism comes from the observation that while the tests measure a little of what students learn in academic courses they measure nothing at all of what they learn in most general and all vocational courses. By their design, therefore, they are missing much of what high schools intend to teach. The general problem of the lack of relationship between the tests and the curriculum is manifested by the very small gains between 10th and 12th grades represented either absolutely (raw score) or relatively (percent of standard deviation of 10th grade scores).

A second problem has to do with the nature of the concepts assessed by the HSB tests. The multiple choice format, the short length of time allocated for the testing and the survey of knowledge nature of the tests reduce the chances for the measures to assess critical thinking or higher order thinking skills. Work is going on around the nation in the area of assessing higher order skills. Fredericksen's work at ETS, the ETS GRE Analytical Score efforts, and Sternberg's work for the state of Connecticut are three examples. In addition there is a lot of exploration of ways of using the computer to create testing environments that assess more than the basic skills.

A third problem is the lack of correspondence of the HSB tests with the IEA, the National Assessment or state assessments. There is, at best, scattered coordination -- the HSB with the NLS72, the National Assessment and some state assessments and, I gather the IEA math, are examples. Someone should systematically set out the inter-relationships among the existing surveys and examine opportunities in the future so that maximum coordination (without sacrificing too much quality) could be achieved.

I recognize the difficulty in solving any one of these three problems much less all three. For one thing, NCES does not have control over either the NAEP or the IEA. My sense, however, is that there is a lot going on in the field of testing that is important, particularly in the theory and development of ways to measure complex thinking skills. There are also advances in the sampling and design areas (eg. BIB spiraling) and in the use of computers to pose complex problems and to tailor tests. Given this ferment it would be well worth NCES's time to get some people thinking about ways of ameliorating the three problems indicated above as well as other test related issues. A few papers and a couple of small conferences that include the NAEP and the IEA folks and a number of the leaders of state assessment efforts might lead to some suggestions that would vastly improve on the measures that are currently used.

3). Work with various parties to build an indicator system. This issue is already on the NCES agenda. I want to reinforce it. Over the next few years NCES should take the lead in thinking about and implementing new and better ways of assessing the health of the nation's educational system. Apart from it being part of the mandate of the agency the time is ripe. There are a number of key roles that NCES should play. NCES should be actively coordinating with the Chiefs, the National Academy of Sciences, the GAO, the NSF, the NIE testing and state and local policy centers and all of the others who now have an interest in federal, state and local level education indicators. Second, while the generic concerns of NCES are allied with those of all of the other actors (high quality data, minimize burden etc.) there are some specific federal concerns that NCES should attend to -- these have to do with indicators to assess the equality of the nation's educational system and the relationship of the system to the nation's productivity and to the nation's defense. A third area where NCES might take leadership is in the development of new strategies for developing key indicators such as ones assessing curriculum or teacher quality.

4). Establish an international program. The quality of data comparing the resources, organization, intentions and outcomes of the various advanced nations of the world is very poor. The IEA surveys, which tragically are state of the art in this area, suffer from lack of funding, proper coordination internationally and, in the US, from a lack of connection with many of the established ways of insuring adequate data collection. This combination of problems has led to erratic schedules for data collection, very poor response rates for US samples, and little use of the IEA data by the US policy system. One way of beginning to ameliorate these problems would be to have NCES assume responsiblity for coordinating US involvement in IEA activities.

Other than the IEA there are only scattered sets of international comparisons which involve the US. Some are carefully carried out though they have only limited generalizablity because of sample breadth (for example, Harold Stevenson's work) -- others suffer from major problems (the Texas newspaper's study). I don't believe that NCES should have much to do with these entrepreneurial efforts except to keep track of them.

There might also be opportunities for cooperation with OECD or with other nations, singly or in groups. My instinct is that there ought to be a mechanism within NCES to seek out these chances -- this should not be institutionalized in a position, however. The Director of NIE used to sit on the CERI-OECD board. This might be an avenue for coordination.

5). Track students from the middle school years through high school and beyond. Analyses of data from HSB have demonstrated the importance of tracking students while they are in high school to document the effects of schools on student achievement and attitudes. Even if the results of cross-sectional analyses on 12th grade outcomes are corroborated by longitudinal analyses of 10th to 12th grade gains the added confidence in the results that comes though the longitudinal addition of earlier grade test scores and attitudes is worth the investment. And, occasionally, it turns out that longitudinal studies provide a very different set of findings than did cross-sectional data. Beyond that, longitudinal data collected in 10th through 12th grade in HSB, in contrast to the data collected only in 12th grade and beyond in the NLS, allowed researchers to begin to examine nationally such issues as students' reasons for dropping out and the short and long term effects of dropping out.

Unfortunately, 10th grade is not early enough to fully examine the effects of high schools on student achievement and attitudes or the reasons of students for and consequences of dropping out, among other issues. By 10th grade many students, particularily poor students in inner cities, have already either dropped out officially or stopped attending school. According to HSB data, for example, 16% of the nation's students drop out before graduation from high school -- most other national estimates which include dropping out prior to 10th grade place the percentage of dropouts as 25-28%. The exclusion of upwards of 40% of the nation's dropouts from the HSB analyses and, in particular, those dropouts who left school prior to 10th grade, suggests that profiles, analyses and policy conclusions developed from the existing HSB data may be misleading if they are interpreted as representing the entire population of dropouts.

Our ability to analyze and understand other high school processes is also limited by studies which gather initial data on 10th graders. The phenonema of tracking and ability grouping is well underway prior to 10th grade in almost all high schools--the determinants of these assignment practices are operating by 7th and 8th grade. The same arguement applies to the comparison of public and private high schools--most of the selection issues have been settled long ago by the time students reach the 10th grade. Still another area of study that is limited by studies which begin in 10th grade have to do with determinants of patterns of participation in extra-curricular activities and of work outside of school. The general point is that inferences about the effects of high schools on students are necessarily limited if analysts do not have data on students that preceeds the entrance of the students into high school.

A related substantive arguement has to do with our collective lack of knowledge about the experiences of students in the middle school years -- grades 6th through 9th. With the exception of the IEA studies(which are very erratic) and the 13 year old samples from NAEP (also erratic, cross-sectional and only recently containing much school and teacher data) we have no systematic data on the experiences, quality of schooling and achievements of middle level school students. If the new high school longitudinal survey were started in the 7th or 8th grade we would gain considerable data about the middle school years, the transition between middle and high school and get a more complete picture of the high school than we had previously.

There are substantial issues having to do with cost, time, and comparability of data sets which would have to be resolved for a program of research like this to be undertaken. Major among these are the various routes that students take from middle level to high schools and the costs of tracking students. My sense is that there are a variety of strategies that could be developed to gain information about efficient ways of collecting middle to high school data -- maybe a small working conference would be useful to do some preliminary planning on this issue.

6). Consider establishing a systematic research program coordinated with Health, Labor, the Census, NSF and HUD. This proposal stems from two motives. First, I am sure that there are inter-governmental coordination committees that meet every two months to share information to help all interested parties in the government stay abreast of the latest and most effective methods of data gathering and analysis. I am equally sure that although these committees are sometimes very effective that a lot more coordination could be done. Second, I am confident that there are a substantial number of ideas in the field about ways of improving data gathering and analysis strategies that could use some stimulation and direction and could provide great savings to the federal government in the very near future. On the data analysis front I have in mind as examples the matching strategies recently examined by Rubin and others and the modeling estimation techniques proposed by Heckman. On the sampling and data collection side the BIB spiraling approach developed by ETS to collect NAEP data is a recent example. Strategies for collecting indicators represents another area that requires systematic research.

In my experience one of the best ways to get agencies to share expertise is to involve them in a collective activity. Of course the activity has to be in an area of concern. And it has to require some committment of resources from each agency. A joint research program among the primary data gathering offices of each of five or six agencies such as Labor, the Census, HUD, the Health statistics office and NSF could be carried out with a modest contribution (say $200K) from each agency for each year.

This would put $1 million in the field to directly address problems that would make the efforts of the agencies more effective and efficient. I suspect it would also lead to greater communication among agencies in data collection and analysis strategies. It might also lead to a greater understanding among the various agencies about the nature and content of data collected by other parts of the government.

7). Consider carefully how to coordinate research and statistics within the Department of Education. Over the 13 years of NIE's existence there has been little systematic cooperation between it and NCES. In the past six months this may have changed. At first blush my sense is that greater cooperation between the agencies and, indeed, systematic coordination of research funding and the gathering and analysis of statistics would be of great benefit and little cost. Such coordination might lead to combining the functions under a single Director. As far as I know, however, the research and statistics gathering and analysis functions are separated in other departments of the federal government. This may be due to the fact that the constituent communities are different, each wanting their own small agency. Or it might be due to a desire to separate research from the on-going policy orientation that the statistics agencies often have. Neither of these reasons appear compelling to me but there may be other reasons to keep research and statistics at arms length from each other. With the NIE being reauthorized and organizational changes to OERI being proposed internally it seems to be a proper time to think through this issue.

END

ISSUES AND CONSIDERATIONS FOR A TEN-YEAR PROGRAM ON ELEMENTARY AND SECONDARY SCHOOL DATA COLLECTION

by

Gail E. Thomas*

National Academy of Sciences
Office of Scientific and Engineering Personnel
Washington, D.C. 20418

June 21, 1985

---

* Gail Thomas is on leave from Johns Hopkins University's Center for Social Organization of Schools. She is visiting the National Academy of Sciences as a National Research Council fellow.

Introduction

In response to a request by the National Center for Education Statistics, this paper will address specific and general issues regarding data needs and considerations that might improve and better facilitate the collection, use, and interpretation of educational data at the elementary and secondary school levels. The issues discussed will range from preschool education to the transition of students from high school to college. While the matriculation of students from high school to the world of work is acknowledged as an important transitional point, research and data issues regarding school-to-work will not be discussed. The reason is that the author's primary experiences and contributions concern the process of schooling and the matriculation and transition of students at various points in the educational attainment process.

Seven specific issues will be discussed. These issues will address data needs and measures based on the students and schools as the units of analyses. Most educational data collection activities at the elementary, secondary, and postsecondary levels have concentrated largely on obtaining student-level data (i.e., student achievement, aspirations and expectations). However, there is a critical need to better understand the process of schooling and other structural components of elementary and secondary education that are associated with and influence various student outcomes. Thus, in focusing on both students and schools, the need for improving past and present data collection activities and suggestions for new data will be discussed. Also, rather than describe specific questionnaire items and measures, the discussion will raise questions and address issues around which measures can be developed. These issues are as follows: preschooling and early education, elementary and secondary school tracking, student attrition, quality of schools survey, teacher and parent surveys, assessing student attitudes toward schools and learning, and the transition of students from high school to college.

Preschooling and Early Education

An important issue regarding early childhood education concerns the nature and effects of preschool education (both formal and informal). Several questions should be raised and addressed by data in this regard. First, what percentage of American children do not participate in preschool education, and why don't they participate? What is the composition by race, sex, and social class of children who do participate in preschool education? Secondly, what is the nature of preschool education, and how structured and centralized are preschool curricula? Thirdly, what are the background characteristics and quality of preschool teachers; and to what extent do they differ from those of elementary and secondary school teachers? Fourth, and most importantly, what are the effects of preschool education on student learning and elementary and secondary school education?

Elementary and Secondary School Tracking

In most national surveys that have included measures on school tracking, data have been obtained at the high school level and have been primarily derived from students, or from school records. These data entail information on the type of high school programs that students are enrolled in (i.e., academic, vocational, general), and the number of credits they have earned in various courses. However, the concept and consequences of tracking extends far beyond these limited measures (Parsons, 1959; Rosenbaum, 1976). For example, the race, sex, and social class stratifications that occur in the classroom are important data that do not exist on a national basis. Also, apart from high school level data, information on tracking patterns in elementary and secondary school are needed.

Thirdly, more detailed data on the type of courses that students take (especially in mathematics and science) will be useful. Studies have found that the type of mathematics and science that high school students take is as equally critical for their access to college (and to various college majors) as the number of mathematics and science courses that they take (Sells, 1976; Berryman, 1983; Thomas, 1984). Finally, institutional or school level data are needed on how students are assigned to various classes and special programs (i.e., special education, and gifted and talented programs) and on the characteristics and qualifications of the teachers that are assigned to these classes and programs.

Student Attrition

Obtaining more accurate and extensive data on student attrition patterns is presently a major data need. In describing the lack of consistency between data collection agencies that report student attrition rates, Cooke and Ginsburg (1985) noted that U.S. Census and

NCES estimates of high school drop-out rates vary as much as 50 percent. A major reason is the different procedures that these agencies employ in measuring attrition. Thus, more coordination among agencies is necessary to resolve or at least to reduce these inconsistencies.

Other important needs regarding student attrition research entail (1) having the necessary data to document attrition patterns much earlier in the educational attainment process (i.e., prior to high school or at the elementary and junior high school level) and (2) obtaining more detailed information on student drop-outs at various educational points. For example, such data should include the race, sex, and social class background of student drop-outs, their records of academic achievement, their attitudes toward school, the reasons that they dropped out of school, and whether they expect to resume their schooling. Also, better data on student re-entry patterns at various points in the schooling process are needed.

Quality of Schools Survey

In recent public documents, it has been argued that the quality of American public education and, subsequently, the quality of its students have declined (National Commission on Excellence in Education, 1983; Coleman, Hoffer, and Kilgore, 1982). However, more detailed and comprehensive data on American public schools are needed to better assess these claims. Therefore, the National Center for Education Statistics should seriously consider undertaking a study of the nation's elementary and secondary public schools. In doing so, several issues might be addressed. For example, one important issue concerns differences between schools not only in the various states, but also between schools within any given state. An assessment should be made both of the uses of standardized tests in elementary and secondary schools and of the differences and similarities in grade promotion and high school graduation requirements from one school system to another. Still another issue concerns the nature of school curricula and the learning process in elementary and secondary schools. In commenting on both of these features, the National Commission on Excellence in Education (1983) noted that secondary school curricula in American public schools have become too diffuse and that students are not being taught the higher-order thinking skills necessary for successful problem-solving. Therefore, data on the type of curricula and teaching and learning paradigms that are employed in different schools and classes (i.e., academic, special education, etc.) should be useful.

Data that would permit a comparative assessment of school environments and student achievement in private versus public schools, and in predominantly black versus predominantly white elementary and secondary schools, are also needed. The latter data will facilitate an evaluation of the extent to which school desegregation has increased educational opportunity and attainment for minority students. Also, the internal conditions of desegregated schools in terms of race relations and the extent to which resegregation patterns exist among students within school classrooms and in extracurricular activities need to be examined based on more systematic data at the elementary and secondary school levels. In addition, data on curriculum differences, school facilities, and the characteristics and quality of teachers and students in public and private schools are needed to reassess the findings by Coleman et al (1982) concerning the superiority of private schools over public schools. A more useful inquiry would entail identifying the elements of private school education that might be useful in improving public education.

Teacher and Parent Surveys

Another important aspect of school quality is teacher quality. Thus, given the dearth of data available on teacher quality and teacher performance, a separate survey on teachers (both public and private) needs to be conducted. It has been recently reported that too many public elementary and secondary school teachers are being recruited from the bottom quarter of high school and college graduating classes and that there is a severe shortage in the number and quality of high school mathematics and science teachers (National Commission on Excellence in Education, 1983). These claims need to be better validated by more detailed and systematic data. In addition, information on teacher salary, teacher attitudes, and teacher motivation are needed.

Data about the parents of elementary, junior, and senior high school students are also needed. These data should assess the extent of parental involvement in schools: the values, attitudes, and extent of familiarity that parents have about schools; and parents' perceptions of the type of education that their children are or are not obtaining. National data on these issues are not available. To assure the participation of low-income families in collecting these data, survey efforts should be augmented by home interviews of a selected sample of these families.

Assessing Student Attitudes Toward Schools and Learning

Data on students' attitudes about schools and their education are very limited and critically needed. In commenting on the importance of such data, members of the Study Group on the Conditions of Excellence in American Higher Education (1984) noted the following:

> An institution that regularly seeks its students' views about the quality of their education experience is manifesting a very different set of values from an institution that makes no such inquiries once the student matriculates. If the only subjects on which we call for student opinion are extracurricular activities, athletics, and food service, we leave the impression that we do not value students as people capable of thinking seriously about their education. (p. 61)

Thus, more detailed and longitudinal student attitudinal data that extends beyond students' self-concept and their educational and occupational expectations are needed. This information should be obtained from high school, as well as junior high school, students who should also be capable of providing useful information about their education.

The Transition of Students from High School to College

The transition of students from high school to college is a critical point in the educational attainment process (Thomas, Alexander and Eckland, 1979). Present surveys sponsored by NCES (i.e., the National Longitudinal Survey of High School Seniors-1972; High School and Beyond) permit rough estimates of student transition rates from high school to college (i.e., two-year vs. four-year; predominantly black vs. predominantly white). However, additional and/or better data on the selectivity of the postsecondary institutions that students attend, the type of financial aid that potential college students receive, and the reasons that students select various types of colleges are needed.

Also, more extensive data on present trends in higher education cost and student finance of higher education are needed. It has been recently reported, for example, that students are increasingly relying on loans rather than grants and scholarships for their postsecondary education (Miller, 1985) and that educational indebtedness is a primary reason that many disadvantaged minority students are not presently attending four-year colleges and graduate and professional schools (Flamer, Horch, and Davis, 1982). Thus, national and state data that would permit an assessment of trends in the availability and packaging of student aid, as well as the effects of various types of aid on different types of students, should be useful.

Other Issues

Four additional issues regarding the collection of future data should be noted in concluding this discussion. The first concerns the need for more accurate and consistent measures of major variables and definitions of minority groups employed in educational surveys. Cook and Ginsburg (1985) have adequately described the severe problem of inconsistency regarding state and national data on student high school attrition. Similar problems have been reported in studies of college student attrition (Panos and Astin, 1967; Pantages and Creedon, 1978). Also, more consistency and coordination among data collection agencies on the definition, aggregation, and disaggregation of minority groups are needed.

High school and college attrition rates and other measures of educational attainment differ not only because of different measurement strategies employed by different agencies, but also because of the manner in which members of the various groups being measured are categorized and defined. This is especially true regarding blacks, Asians, and Hispanics—whose educational status and attainment may differ substantially depending upon their cultural and social backgrounds and their country of origin. Thus, more detailed and less gross distinctions (i.e., minority vs. nonminority) should be employed in the recording, collection, and reporting of statistical data on minorities.

Secondly, greater efforts to coordinate data collection activities among state and federal agencies and among researchers are needed. This should assist both in reducing inconsistencies in the conceptualization and operationalization of measures and in reducing unnecessary duplication in collecting data. Thirdly, NCES should be commended for consulting with and obtaining input from data users and researchers during the initial stage of its proposed ten-year program. In fact, these efforts should be continued throughout the various stages of the program.

Fourth and finally, a major priority should be established by NCES to continue and extend longitudinal data collection efforts. Such data will continue to be critical in studying schooling processes and student achievement and attainment over time. NCES' National Longitudinal Survey (NLS) of high school seniors was a massive and important investment that should be continued and extended by subsequent follow-ups. One immediate and important need that a subsequent follow-up of the NLS might serve entails a study of the process and experiences of students in U.S. graduate and professional schools. Also, applying and consistently maintaining a longitudinal approach to the current program that NCES is proposing (especially given a focus on pre-high school and early education) should prove extremely valuable in filling many of the existing voids in educational policy and research.

References

Berryman, Sue E.
1983 Who Will Do Science? New York: The Rockefeller Foundation.

Coleman, James S., Thomas Hoffer, and Sally Kilgore
1982 High School Achievement. New York: Basic Books.

Cooke, Charles and Alan Ginsburg
1985 The sorry state of education statistics. Basic Education 29: 3-8.

Flamer, Herber J., Dwight H. Horch
1985 Talented and Needy Graduate and Professional Students. New Jersey: Educational Testing Service.

Miller, Scott E.
1985 Student and Parent Loans: A Growing Reliance. Washington, D.C.: American Council on Education.

Panos, R. G. and A. Astin
1967 Attrition among college students. American Educational Research Journal 5: 57-72.

Pantages, T. J. and C. F. Creedon
1979 Studies of college attrition 1950-1975. Review of Educational Research 48: 49-101.

Parsons, Talcott
1959 The school class as a social system. Harvard Educational Review 29: 318-329.

Rosenbaum, J. E.
1976 Making Inequality. New York: John Wiley.

Sells, Lucy W.
1976 The Mathematics Filter and the Education of Women and Minorities. Paper presented at the annual meeting of the American Association for the Advancement of Science. Boston, Massachusetts.

The National Commission on Excellence in Education
1983 A Nation at Risk. Washington, D.C.: U.S. Department of Education.

Thomas, G. E., K. Alexander, and B. K. Eckland
1979 Access to higher education: The importance of race, sex, social class and academic credentials. School Review 87: 133-156.

Thomas, G. E.
1984 Determinants and Motivations Underlying the College Major Choice of Race and Sex Groups. Baltimore, Maryland: Center for Social Organization of Schools.

COMMENTS ON AN ELEMENTARY AND SECONDARY EDUCATION DATA PROGRAM

Brenda J. Turnbull
Policy Studies Associates

The National Center for Education Statistics deserves credit for undertaking a fundamental redesign of its data program in elementary and secondary education. In this paper, I suggest that NCES use this opportunity to (1) analyze likely uses of the information it collects and disseminates, (2) launch more detailed investigations of educational inputs and processes, (3) take some steps to place findings on educational outcomes in context, and (4) continue to attack the problem of quality control.

First, a word about my vantage point is in order. I am a researcher who collects and analyzes information about state and local implementation of programs for special-needs students and for educational quality. I try to draw from this information the types of conclusions that can help policymakers assess and modify programs. Thus I have a bias toward the information that policy audiences will use. I also like to see detailed information about what goes on in schools, which is not always conveyed by gross measures of the resources put into schools. Finally, because my familiarity with NCES is only that of an occasional user of its data, this paper may recommend some procedures that are already standard practice at the agency. If so, please consider those comments endorsements rather than recommendations.

Addressing the Needs of Information Users

Ideally, the collection of data should be driven by a framework of questions that the data will be used to answer. Working backwards from intended uses through projected analyses to the specification of data elements and methods of collecting them would result in an efficient and practical program. In reality, NCES has to live with uncertainty about the questions that will be posed. However, there ~~are some~~ ways of reducing this uncertainty, in addition to soliciting advice through commissioned papers and hearings. Analysis of information needs could include the following approaches:

- Identifying and classifying the concerns about elementary and secondary education recently expressed by important groups such as Congress, governors, state legislators, and the public.

- Doing a content analysis of recent reports and reform proposals, looking for both the indicators that have convinced their authors of the poor health of the education enterprise and the types of improvements in process and outcomes that they believe their recommendations could produce.

- Drawing analogies in education to the statistics and indexes that are used in other fields, such as the gross national product, measures of housing starts, and the like.

Classifying the concerns of various constituencies

One guide to the future interests of policymakers and the public is what they have said about education in the recent past, since these groups' concerns remain relatively stable over time. The Gallup poll consistently shows that the public is concerned about discipline in their local schools, for example. The members of Congress for whom the educational opportunities of special-needs pupils have been an important concern are not abandoning their interest in these pupils. Governors and state legislatures will remain concerned that their states offer employers a skilled workforce.

All these concerns can be translated into indicators that are worth collecting on a national scale. Citizens' worries about discipline suggest that it would be useful to have data on the frequency of various types of incidents in schools over time. Because the education of special-needs groups remains an issue in federal policy, data on educational resources, processes, and outcomes should be broken down by type of pupils wherever possible. The skills of entry-level workers, besides being a state concern, deserve some analysis on a national scale.

Even some concerns that are primarily local can provide clues to data that are worth collecting nationally. Local administrators routinely collect and often use data on student attendance, the popularity of particular course offerings, and turnover among teachers and aides. Collecting some sort of data on matters such as these might be feasible and worthwhile on a national scale since these topics reflect legitimate concerns about the workings of school systems.

Analyzing the indicators cited in recent reports

What evidence has recently convinced the public and the media that education in the United States is in sad shape (to the extent that the flurry of enthusiasm for reform has been based on evidence)? An analysis of the indicators in which the authors of the reform reports have found cause for gloom and those in which they urge improvement could provide another source of ideas for data collection. The report authors, along with the press and television, succeeded in making a case to the country. The types of evidence they used might reasonably be expected to make sense for future information users, too, if collected on a continuing basis.

This does not mean that every fact or impression cited in the rush to condemn and uplift the schools deserves to be formalized as part of the Common Core of Data, naturally. One of NCES's responsibilities should be to place the more sensational findings in context. For example, comparisons of achievement in the United States and other nations ought to be viewed alongside other measures that help to balance and explain them, such as data on the proportions of children in various countries who attend college or reach the age of 18 with the qualifications for college admission. The point is that if international comparisons (and other types of evidence on the health of our schools) have persuasive force for the public, then they ought to be drawn and presented in a way that is as technically defensible as possible.

Drawing analogies to indicators in other fields

Although statistics on education are not going to be watched as closely as those on the economy, some ways of collecting and reporting economic data may provide useful models for NCES. For example, the index of leading economic indicators and data on housing starts both provide clues to the health of the economy, and they are especially useful because they provide current data that tend to predict future developments. A concerted effort by NCES to develop and publicize analogous data could serve the worthwhile purpose of giving observers of the education system something to look at besides test scores.

Data exist on the number and characteristics of college freshmen considering careers in teaching, for example. These data could be publicized as an annual signal about coming changes in the teaching force. It might also be possible to aggregate a number of measures of educational achievement and attainment into an index like the gross national product. Such an index would oversimplify matters, of course, but so do the economic indexes in widespread use.

In summary, measures of educational trends that are intuitively easy to understand, that capture changes in several important dimensions at once, or that tend to predict future changes could meet some information needs of the public and policymakers. If they were released with some fanfare and accompanied by clear explanations, they might also help educate people about what goes into educational quality. Analysis of various groups' concerns about education and the data they have found persuasive in the past can also help guide the development of such measures.

Describing Educational Inputs and Processes More Fully

Alfred North Whitehead said, "Seek simplicity, and distrust it." Large-scale data collection is especially useful when it produces descriptive data, and simple descriptions of such things as the demographic characteristics of students and teachers are very useful for researchers, policymakers, and the public. On the other hand, broadly based surveys aimed at describing what goes on in schools run a serious risk of producing bad information. My experience in schools suggests that resources often go unused or are used in unexpected ways. My recommendations, therefore, are twofold:

- o Do not underestimate the value of the simplest demographic statistics.
- o In the effort to describe more complex educational resources and processes, probe beneath the surface for detailed data on exactly who does what with what resources.

The value of simple statistics

NCES publications already contain the types of simple, factual data that are indispensable in compiling a portrait of the education system. Data on the characteristics of students, teachers, schools, and school systems provide a sense of trends in education and important background variables for analysis of developments at the national and state levels. Time-series data represent a unique resource for research and policy analysis, and the federal government is particularly well qualified, by virtue of its centrality and visibility, to collect such data.

The value of demographic and other descriptive data increases greatly when the data are collected and summarized in a consistent way from year to year. When a measure changes, the trends it is intended to capture can be exaggerated or masked. Therefore, even when there seems to be good reason to tinker with a question or an index, the potential improvement should be carefully weighed against the loss of comparability. Stability in measures should win out in most cases.

The urge to improve on existing data can find a useful outlet in extending data collection to new areas. Early-childhood education and other child-care settings are an example of such an area, where new developments are occurring (and policymakers may eventually see a need to catch up).

Finally, when descriptive data have been collected, they should be published as rapidly as possible. Because delays in publication reduce the usefulness of data, the Department of

Education should expedite whatever steps are needed for clearance and quality control.

Measuring resource use and program characteristics

Over the past fifteen years or so, evaluations and other studies of educational programs have grown far more sophisticated because of the recognition that a program may exist on paper and yet make limited or unpredictable differences in classroom practice. This problem of implementation deserves analysis as it applies to the work of NCES. Although I argued above for some oversimplified indicators that will meet the public's information preferences, I also believe that sometimes a simple measure of educational resources is worse than no measure. Some statistics give a misleading picture of what goes on in schools, and for selected issues the extra care and expense of gathering accurate, detailed information will be worthwhile.

A topic that has received recent attention is that of microcomputer purchases and use in schools. Local news features often report how many schools "have computers" or describe parents' fund-raising efforts to buy one or two microcomputers for a school. The administrators seem to have correctly judged that they would receive as much public credit for buying one machine as fifty. NCES should not fall into this trap, however. The 1984 edition of The Condition of Education usefully reports on how students use microcomputers and how many minutes the average student user spends with them each week. With the addition of some numbers indicating how many students are users, this would be a good summary of the key information about microcomputer use. It might be even more useful if it were supplemented with some data--even anecdotal data--illustrating the types of use summarized. What is covered in courses in computer literacy, for example?

Innovative technology is not the only area in which the details of resource use deserve scrutiny. The number of teachers reportedly present in a district or school may not translate directly into measures of class size or even of the actual teaching force. Not only do some teachers move around the building teaching lessons in music or art, but some of them never instruct students. They are resource teachers, assistant principals (in name or in fact), or managers of categorical programs. I met several of these nonteaching teachers when I did fieldwork on Title I of the Elementary and Secondary Education Act. In setting up the interviews, I had asked to see Title I teachers; several principals arranged for me to meet with people who held teaching certificates and were counted on the building roster as teachers but never worked with students other than to test them.

In addition to gathering more detailed data about the use of particular resources in schools, NCES might probe more deeply into descriptions of curriculum. At the secondary level, we need not only surveys of the courses that students take but also some data (probably from teachers) on the topics that the courses cover, the relative emphasis given to these topics, and the texts used. Elementary teachers can furnish similar information on the content they try to cover and the texts they use. Because such data can be especially useful in conjunction with measures of student achievement, the National Assessment of Educational Progress (NAEP) is one logical vehicle for collecting them.

At a further level of detail, data on instructional processes would be useful if some reasonable level of accuracy is possible. Again, teacher surveys could indicate how time is used during the school day or the class period, how many interruptions occur and what these interruptions are, what is taught in small groups and in whole-class instruction, what disciplinary incidents occur during a typical day, and so on. Naturally, better questions will produce better data, and teachers should take a major role in helping to refine any such surveys.

The analysis and presentation of detailed data on school programs must differ in some ways from the analysis and presentation of the demographic and fiscal data that NCES has most commonly collected. Nationwide aggregates and averages may not mean very much. For example, the data in Indicators of Education Status and Trends on the presence of remedial courses in college undoubtedly obscure massive differences in the level of remediation needed in colleges with different levels of selectivity. Because of the extent of variation in such courses, there can be legitimate debate about whether the nationwide enrollment figures mean anything.

Reports on course content and instructional processes should contain information about the amount of variation found as well as the central tendencies. Breakdowns by type of district (large city, small city, suburban, rural) and, if possible, by state would also be interesting. Finally, data on the educational experiences of particular student groups (for example, girls, boys, students with handicaps, students attending high-poverty schools, low achievers, and those whose native language is not English) could inform the assessment and formulation of federal policy for these groups.

In summary, I would argue for the collection of detailed data on schools' resource use, curriculum content, and instructional processes. The expense of doing such data collection well means that it must be restricted to only a few topics, but even data on selected details of educational practice would be useful. They would supplement the bread and butter of data collection on the simpler characteristics of students.

teachers, and school systems--which should continue with as much consistency over time as possible.

Measuring and Reporting Outcomes

Undoubtedly, most of the attention that citizens and policymakers give to education statistics will continue to center on students' test scores. Perhaps, though, NCES will be able to make marginal contributions to a more complete picture of the outcomes of education. It can do the following:

- Accompanying reports of test scores with more description of what was tested.
- Giving publicity to outcomes such as attainment.
- Seeking, using, and explaining measures of higher-order skills.

Describing the tests

Reports of students' scores on tests should be closely accompanied by descriptions of what the tests covered. The displays of NAEP results in The Condition of Education indicate the general categories of knowledge tested, but putting somewhat more detail on the pages that contain these summary tables and graphs would be even better. Interested readers should also have ready access to comparisons between the NAEP tests' content and the content of the standardized tests that school systems commonly use.

Ideally, the media and state and local school systems would then follow the federal lead and describe what high and low scores mean. In any case, NCES should do so.

Measuring other outcomes

Education professionals know that test scores do not summarize all the interesting outcomes of schooling. NCES and the rest of the Department of Education should not only continue to measure other outcomes but also emphasize them whenever possible. Good measures of dropout rates--and of the number of people who obtain general equivalency degrees--are examples of such outcome measures. The extent of voting among recent high-school graduates is another interesting statistic to have.

Emphasizing higher-order skills

This recommendation reflects my sense that schools are increasingly teaching to the tests devised by test publishers and state governments, that these tests tend to measure rather simple skills such as decoding and arithmetic operations, and that at some point there may be a backlash from education professionals or even employers who notice that other outcomes are being neglected. Defining higher-order skills is hard, and measuring them is harder, but the attempt would be worthwhile.

Because I have very little technical knowledge about measurement, my recommendations in this area are fairly general. However, as a would-be consumer of data on student outcomes, I can advocate that NCES continue and perhaps intensify its efforts to place test scores in context.

Quality Control

Because all statistics imperfectly represent reality, an agency that collects and disseminates statistics can never do too much to identify their weaknesses, correct the weaknesses it can correct, and explain the ones it cannot. NCES has a further challenge stemming from its reliance on diverse state education agencies (SEAs) to collect and report standard data. Quality control takes different forms for the collection of data from national samples and the aggregation of state-collected data:

- Surveys of national samples will provide the highest-quality data on topics where terms have varying definitions, such as "dropouts," and their quality can be improved if studies of response bias are built in.

- NCES should not hesitate to take a strong stand with SEAs on quality and consistency in the data they provide to the federal level.

Surveys of national samples

National surveys like High School and Beyond perform a unique service to research and policymaking in education. Although expensive, they provide relatively trustworthy data on detailed topics and permit the analysis of a variety of educational issues at the level of the student, the school, or the community. Where inconsistencies arise in state-level data because states define terms in different ways, national surveys using standard definitions can help resolve the resulting questions.

National studies present issues of response bias that are manageable. For example, we know that students tend to have erroneous ideas of their parents' income levels and that household surveys provide better data on this topic. More could be done to identify other sources of faulty data in national surveys so that alternative means of data collection (such as different respondents or differently worded questions) can be tried.

Aggregation of state data

Some reliance on SEAs to provide data to NCES is sensible and even unavoidable, given the relative amounts of resources available for data collection in education at the state and federal levels. However, the problems resulting from poor quality control in states or inconsistencies across states are not entirely unavoidable. NCES should be willing to take a firm stand on quality control.

The overall objective should be for SEAs to recognize that they are participating in a process that is centrally controlled in order to meet national information needs in a technically defensible way. For example, NCES should check the data received from SEAs against last year's data, look for internal consistency in each state's data, and examine how terms are defined across states. Any questions or problems should be raised with the state, and all states should recognize that they may have to revise their initial submissions. Other quality-control steps will probably be necessary as well.

In this federal system, Washington defers to state authority on many issues. Oddities of data collection and analysis should not be among them, however. I believe that, although the states rightly guard their prerogative to collect their own data in their own ways, they are probably more willing than NCES thinks to go along with a national data program that maintains high standards of professional quality. State officials may grumble about the imposition of tighter central standards, but some friction is inevitable in any intergovernmental endeavor, including the present system of more loosely controlled reporting.

The Council of Chief State School Officers can be a major ally in a nationwide effort to correct the flaws in current data aggregation. Other state-based organizations, such as the National Governors' Association, the National Conference of State Legislatures, and the National Association of State Boards of Education, might also play a role, since their members also have a stake in good national data on education. The current process of generating papers, summarizing their recommendations, and holding hearings can lay the groundwork for specifying action steps that SEAs could take. If SEAs and others are not already engaged in developing such steps, they should begin soon.

In summary, although NCES lacks the political or legal power to compel states to do much about their collection and analysis of data, it can probably muster the professional authority to improve the data it collects from states. Together with the further refinement of national surveys, this stance on quality control should go far towards correcting the weaknesses currently visible in national education statistics.

Conclusion: Setting Priorities

Internal deliberations at NCES as well as this process of public comment will undoubtedly generate many good ideas. Deciding where to concentrate the agency's efforts will be the next problem. Although NCES staff will soon immerse themselves in numerous detailed design issues, a focus on a few overriding principles will be in order. I would suggest that the highest priority be placed on strengthening the credibility of the data through such means as quality control and consistency in data series. Next in importance would be the principle of addressing information users' concerns about the education system, as best these can be inferred in advance. To the extent possible, NCES might then expand its data collection in areas such as the detailed description of instructional resources and processes. Whatever principles are chosen as the key ones, the important thing is to concentrate on a manageable set of priorities in this ambitious redesign effort.

Educational Data Needs for the Balance of the 20th Century: Some Perspectives on the Emerging Environmental Context

(A Paper Prepared for the National Center for Education Statistics)

Michael D. Usdan
Institute for Educational Leadership
June, 1985

This paper is written from the perspective of a generalist who is concerned about the whole panoply of social, economic and political issues which impact elementary and secondary education. The author is not a data analyst, nor does he have particular expertise in the whole area of education data collection. The perspectives in this paper will be those of a generalist who believes that the data historically collected in education have been too isolated from other information sources. My contention is that educational data in the future should not only be derived from schools and other formal educational institutions, but also from other deliverers of educational or training services such as the private sector, the military, voluntary associations and the countless other organizations and agencies which provide education and training services.

Education must be defined more generically to consist of more than just schools. Outcome measures relating to student performance should be stressed and significant non-school centered social and economic information from a broad range of data sources both governmental and non-governmental in origin must be employed in efforts to explain student achievement. Definitions of students and analyses of the learning process should be broadened to include recipients of educational services in a whole range of non-school settings. Data must be packaged efficiently and more concisely and presented in ways that are easily understood by the lay public and not just by education specialists.

As the recent reform movement indicates, political and civic leaders, as well as the general public, are willing to support increased funding for schools if they are convinced that the dollars expended are generating positive results. In other words, there will be more dollars for education only if there is more education for the dollars. This political fact of life augurs an era of accountability in which schools in the increasingly competitive struggle for resources will have to justify what they do to a much great extent on a cost-benefit basis. Unless the current weaknesses of national education statistics are remedied and existing inconsistencies and inaccuracies ameliorated, justification of increased expenditures will not be supportible or politically viable and the necessary fiscal support will not be forthcoming in an aging nation in which fewer than 30% of the adults have their own children enrolled in elementary and secondary schools.

The reform movement and the heightened interest of Governors, civic elites and influential business leaders will not last indefinitely and certainly will not be sustained without appropriate and supportible data that can indicate which reforms work and which do not. The resources will not be available to support all of the multiple reforms being enacted in states throughout the country. We need to develop data that are qualitatively better and comparable within and between states. We also must build better indicators that can help us determine more definitively the elements that strengthen student achievement or performance. The timing of the NCES efforts to reassess in a comprehensive manner its data gathering processes could not be more propitious because of the unique public policy saliency now being enjoyed by elementary and secondary education because of the current reform movement.

Let us now look at the larger societal context in which elementary and secondary education data issues must be addressed. Broader human capital and adult learning issues also must be addressed in a rapidly changing and increasingly interdependent social, political and economic environment.

What are some of the major issues which will so markedly influence our future and have such consequential ramifications for all of us? My favorite philosopher is Pogo who says that we are surrounded by insurmountable opportunities. I believe that this is indeed the case in education where the recent spate of nationally publicized reports, as mentioned earlier, has given the enterprise positive "front-burner" public visibility in ways that have not been experienced since the Sputnik era almost three decades ago. If one accepts the old adage that invisible politics is poor politics, then the existing window of opportunity should be capitalized upon, and education at all levels should remain part of the "warp and woof" of major public policy issues and concerns. Indeed, some skeptics are fearful that the educational balloon may be deflated as quickly as it was inflated once the political rhetoric and the freshness of the recent highly publicized national reports and recently enacted state reforms have abated. A growing number of the nation's most prestigious business and governmental leaders are now espousing the importance of education and human resource development to the nation's economy and comprehensive reform legislation has been enacted in states like Texas, South Carolina, Mississippi, California, and Florida among others. Governors like Hunt of North Carolina, Dupont of Delaware, Robb of Virginia, Alexander of

Tennessee, Riley of South Carolina, Winters of Mississippi and others have been in the forefront of political and economic efforts to strengthen the quality of and support base for education. Business leaders from major corporations such as IBM, Control Data, American Can, Hewlett-Packard, the Bank of America, Dow Chemical, Proctor and Gamble, and numerous other representatives from the private sector throughout the country are becoming engaged in education and related human resource development issues in unprecedented and varied ways.

Major national business organizations such as the Conference Board, the Committee for Economic Development, the Chamber of Commerce, the National Alliance of Business, the Business Roundtable, and their counterparts at the state and local levels are undertaking education-related activities or discussing the significance of education issues with unusual interest and even zeal in a number of cases. Education is being viewed more commonly now not as a consumer of resources but as a necessary investment in human capital and the nation's economic viability. This interest and support will be sustained, however, only if we have quality data that can be used to demonstrate the cost-benefit payoffs of various instructional programs as manifestd in improved student performance.

I would like to amplify these points by stressing a number of major issues which should help to illuminate the kind of data that will be necessary, leaving the task of defining the specific data to be gathered and procedures to be followed to others more qualified than I am to discuss these matters. I hope that this "environmental scanning" will be helpful to NCES as it hopefully identifies its data collection strategies not in institutional isolation but within the context of complex issues in an changing and interdependent social matrix. These issues transcend education per se and will require diverse information that entails many factors that are not directly related to the school environment.

Among these somewhat interrelated issues are (1) the need to develop and invest in human resources; (2) the implications of changes in the nation's demography; (3) the importance of reconfiguring and redefining education; (4) the impact of technology on educational and economic development; (5) the ramifications of an aging and changing work force; (6) the special and acute problems of urban and minority youth; (7) the need to define and achieve quality at all educational levels; (8) the weaknesses of foreign language instruction; (9) inter-level educational relationships; and (10) the development of an appropriate federal role in the shaping of educational policy.

These ten issues, of course, hardly represent either a comprehensive or all-inclusive litany of major concerns. Some of these issues will be discussed rather extensively, while others will be alluded to briefly or barely mentioned. Collectively, however, they do form the basis for projecting a plausible change-laden future scenario in which education issues will be of great import and policy decisions must be predicated upon richer and more diverse and reliable sources of information.

The Need to Develop and Invest in Human Resources

In the past few years, the importance of education has been discussed in the newer and broader context of its central role in the collaborative efforts that will be necessary if the nation's economic productivity is to be improved, its relative economic decline counteracted, and the erosion of its international competitive position arrested. Education is pivotal to these national concerns as is the growing awareness that education remains the key to maximizing the nation's human capital and economic development. Support for education in the business community, for example, increasingly is being viewed not as mere philanthropy but as sheer self-interest. The economy and its needs, in other words, are the contemporary Sputnik and serve as a lever for the support and reform of the educational enterprise. In essence, the key to education's new salience is the society's need to develop its human capital at all age levels with emphasis not just on traditional school age youngsters and institutions but also on the large numbers of adults who need retraining in a period when the economy is undergoing such wrenching transformations. Education, then, as mentioned earlier, ought to be defined more generically as consisting of much more than formal elementary and secondary schools or colleges serving the younger segments of the population. Data should be gathered and shared between and among the diverse educational delivery systems and the isolation of schools from other providers of education should be ended. Information about successful teaching and learning processes in other environments as well as in schools ought to be analyzed and disseminated reciprocally so that there are constant exchanges between schools and other deliverers of education and training services.

There is a growing appreciation throughout the society that human resources are the dominant factor accounting for growth in national income and that they account for the major share of the nation's total economic output. Recognition is mounting that the current economic challenge is as fundamental as the change from an agrarian to an industrial economy after the Civil War and that education generically must be linked inextricably to economic development and be viewed as an essential investment in the future. *Data on the educational process and ways of improving student achievement should be gathered from all deliverers of education services and not just from traditional schools serving 5 to 17 year-olds*.

The Implications of Changes in the Nation's Demography

Demographic data are of particular importance to the entire society. In a world in which more than 75% of the total population is of color, we are going to have redefine what the words "majority" and "minority" mean. Indeed, more than 25% of our public school population already is of minority origin, and the proportion of Blacks, Hispanics, Asian-Americans and Native-Americans continues to grow. By 1990, 25% of the total population will be minority. In the 1982-83 school year, for example, 46% of the public school enrollment in Texas was minority, 43% was minority in California, 32% in Florida and New York, 33% in Arizona and Maryland, 28% in New Jersey, and 57% in New Mexico.

The Hispanic population is the fastest growing segment of the minority population with approximately two-thirds of the nation's Hispanics being located in only three states (California, Texas, and New York). More than 85% of the Hispanic population lives in only nine states (New Jersey, Florida, Colorado, Illinois, New Mexico, and Arizona in addition to the aforementioned three states). This dramatic skewing of the distribution of the Hispanic population creates a serious problem of demographic illiteracy as most of the country remains blissfully ignorant of the full import of the dramatic growth in the number of Hispanics. Indeed, Hispanics are the youngest population group so additional increases in the Hispanic population are quite predictable.

California, our largest state, by 2000 will be "majority-minority." Twenty-three of twenty-five of our largest school systems already are "majority-minority" and unless we improve the quality of the education and the life chances of these massive cohorts of minority youngsters, the social fabric of our society may well be imperiled.

The demographic phenomenon of "gray power" also will be of increasing significance as the population ages and the number of citizens 65 years of age and older rapidly expands. By 1990, the number of youngsters under 20 will fall below 30% of the nation's population for the first time in history. In fact, in 1983, there were more people 65 years of age or older than teenagers. The implications of these changes are brought forth starkly by an analysis of the ratio of working to retired people. In 1940, 10 citizens worked for every retired person, in 1985 the ratio is 5.3 to 1 and a 4.7 to 1 ratio is projected for 2000. By 2030, a 2.7 to 1 ratio of workers to retired persons is considered possible.

Thus, growing numbers of retired individuals will be dependent upon a younger population, increasingly minority in composition, to sustain the economy in general and to maintain the viability of an already rather vulnerable social security system.

It is estimated that more than 90% of the 1990 workforce and more than 80% of the 2000 workforce already are employed in the market place. Thus, traditional schools serving the young will be dealing with only a small proportion of the workforce, a full one-half of this younger population will be minority youth who will constitute, as noted earlier, a much higher percentage of the next decade's diminished youth cohort. It is projected that by 2000, approximately one-third of the young will be socially and/or economically disadvantaged.

Such demographic data are of singular importance and should be widely disseminated not only to educators but also to citizens throughout the society. These rather startling numbers reflect compellingly why traditional schools will have to broaden their base of political support with less than 30% of the population currrently having youngsters enrolled. Educational leaders can no longer assume the public support they could generate, for example, when 60 or 70% of the population had children in the schools in the rapid growth era after World War II. Schools will have to reach out to cultivate new and broader constituencies if they are to acquire the necessary support as older citizens and other groups make increased demands for services in a period of declining resources. Public schools will have to be viewed as a civic responsibility by the majority of the population which will not have children involved or a direct stake in the educational process. Educators can no longer assume that they have a broadly gauged support base and certainly cannot afford the luxury of pulling the wagons around in a circle and shooting inwardly as has often been the case in the intramural or internecine warfare which has occurred with some frequency over collective bargaining, church-state relationships, funding allocations, desegregation, interlevel responsibilities, and other issues.

The Importance of Reconfiguring and Redefining Education

As part of the efforts to broaden their support base, educational leaders will have to reconfigure and essentially redefine education as constituting more than just traditional schools or formal institutions of learning. The expanded coalition which they will be compelled to build should include preschool children, senior citizens, the numerous older workers who will have to be retrained, and others in the society who will require educational and/or training opportunities in a changing economy. Multi-agency and collaborative human service thrusts within the public sector and growing partnerships between the public and private sectors will become more essential as resources decline. Educators have unique social and institutional penetration into the grassroots of virtually every community. This outreach will be a great asset to them in efforts to reach out politically not only to buttress and expand their own support base but also as they seek to help fill the escalating educational needs of adult groups such as senior citizens, older workers who need retraining, and single and two-career parents with "latchkey" children. *As these collaborative efforts evolve, data sources, of course, would have to be expanded and diversified to accommodate to the needs of a much more heterogeneous student population with more pluralistic learning objectives*.

Our society must recognize to a greater degree than it has thus far the incredibly diverse and extensive array of non-mainstream educational service providers such as industry, labor unions, some 9,000 proprietary schools and colleges, 300 businesses with a site called "college" or "university", the far-flung and extensive military educational system, hospitals which provide their own educational programs, major service agencies with educational programs such as the Girl Scouts and United Way, diverse day care centers, and so on almost *ad infinitum*.

Industry, for example, invests approximately $40 billion in education and training, a total which rivals the investment made in traditional public higher education. Control Data spends hundreds of millions of dollars in its Plato system and innumerable other corporations such as IBM, Xerox, and General Motors are allocating substantial resources to improve the quality of education and training received by their employees. Wang, Rand, Northrop and A.D. Little actually offer degrees under the aegis of their corporate educational activities and Hamburger and Holiday Inn Universities are realities. Such non-traditional post-secondary educational enterprises already are educating well over one-half of the adults in the nation. This is hardly an inconsequential shadow system and traditional educators and other providers of human resource development services must be more cognizant of the collaborative opportunities as well as the potential competitiveness of this somewhat parallel and non-traditional system. More information must be systematically compiled about this important and rapidly expanding private sector based educational system.

The Impact of Technology on Educational and Economic Development

The area of technology confronts the entire society with a host of challenges and opportunities. The unprecedented technological revolution which confronts us will reshape virtually every facet of our lives in varied and unpredictable ways. The world of computers, discs, satellites and so forth is upon us and our young people from all socio-economic backgrounds will be motivated in new ways to partake in the technological era which has dawned.

Fields like science and math, of course, are essential for technological progress and our society's need for well educated citizens in these areas increasingly is being recognized. The crisis in math and science education already is generating significant changes in educational policy which would not have been politically viable just a short time ago. Notions like merit pay, differential salaries and roles for teachers, and the recruitment of part-time teachers from the private sector or the ranks of the retired would not have received serious attention in either state legislatures or the U.S. Congress without the widespread national consensus that our economic development is inextricably linked to the quality of our educational system and that our nation's scientific literacy cannot be permitted to lag any longer if we are to compete in an increasingly interdependent world economy. The saliency of this issue compels us to carefully gather information about the successes and failures of the varied efforts which are being initiated throughout the country to attract more qualified science and math teachers.

The reality is that the tempo of economic and technological change is such that partnerships between the public and private sectors are essential. There simply is no way in which schools or most other public entities can maintain state-of-the-art equipment or keep up with the new technologies. Creative tax and related policies must be developed which facilitate intersector collaboration if we are to meet the international economic challenges from nations like Japan and West Germany. Information ought to be collected and disseminated which focuses upon what works and what doesn't work in fostering cooperation between and within sectors.

The Ramifications of an Aging and Changing Workforce

The changes which are shifting the foundations of our economy obviously have deep ramifications for educators. Dramatic and continuing reductions in the manufacturing or industrial areas will require the development of a workforce better able to perform in the information and service providing realms. We will thrive if we sell our wits - if not our goods - to the world and our major future resource is brainpower.

The challenge educationally is not only to improve basic skills but also to react to the changes in occupational context which will compel more workers to shift from "take/place/lift/put" jobs to work like computer programming which requires listening and reasoning skills as well as adequate backgrounds in fields like math and science. These challenges cannot be met by single sector approaches. As a society, we can no longer afford the luxury of fragmentation or "turfsmanship". The challenges will require institutional collaboration between and among education, business, labor and government. Education and training must be viewed more generically and as a multi-sector and societal responsibility in which we all have a stake and role. The Digital Corporation, for example, allocates from 10 to 20% of the time of its engineers and other highly skilled personnel for in-service training so that they can keep abreast of the phenomenal rate of technological change which gives the schooling of a newly-graduated engineer only a five year period of relevancy. Organized labor also is acutely aware of the continuing need of reeducating and retraining its members. Indeed, recent contracts in the auto industry, for example, explicitly provide time for training or educational activities.

As a society, we can no longer afford parallel and costly human resource development systems which rarely if ever interact and collaborate. We are no longer rich or wealthy enough to afford such costly duplication and fragmentation. Business, for example, is compelled to spend many millions of dollars on remedial education, a responsiblity which quite logically could have been assumed to have been within the purview of the schools. AT&T, for example, which spends almost $2 billion annually for educationally related activities, expends considerable money teaching thousands of employees basic writing and arithmetic skills. Metropolitan Life, like numerous other corporations, devotes a large porportion of its training program to remedial work in the basics. Efforts should be made to collect information about such remedial efforts to see if certain practices can be effective in public institutions.

Despite these problems, the United States has great advantages once we "get our act together" and generate the necessary intersector linkages. We have a strong base and tradition as a creative entrepreneurial society and our education system despite its weaknesses has produced a uniquely well educated and adaptable population. Parenthetically, it is important to point out that efforts by educators to build partnerships with the private sector should not be limited only to large corporations. Approximately 580,000 small businesses are started each year in this country and two-thirds of the new jobs are created in smaller enterprises. Thus, the opportunities for partnerships between education and business in many cases can be more effectively developed at the local or regional level where much of the action if not the rhetoric takes place.

A dominant and recurrent theme articulated in all of the recent reports on education relates to the concern about the quality of the teaching force. There is widespread consensus that any efforts to strengthen education and the development of our human capital must be predicated upon improving the status and intellectual caliber of the classroom teachers who are at the core of the educational enterprise. In recent years with the constraints on governmental spending there has been a drain of talent from the public sector and service fields in general. Teaching without question is not attracting our "best and brightest" young people and more vigorous attempts must be made to replenish and strengthen the ranks of those responsible for developing the future intellectual capital which is so vital to our security and well-being.

The problem is most acute, of course, in areas like math and science where the economic and status rewards of teaching cannot begin to compare with the salaries and recognition available to talented individuals in business and industry. We have finally recognized that we have a crisis of national proportions with approximately only one-half of our math and science teachers appropriately certified. Throughout the country numerous proposals are being made to differentiate roles and salaries, develop merit pay schedules, and recruit part-time instructors from the private sector as efforts intensify to attract and retain talented teachers. A number of school districts have already implemented such plans and recently large states like Florida and California among others have enacted comprehensive reform legislation which has included provisions for merit compensation or differential categories of teachers who would receive extra pay. Data on these efforts should be collected and careful evaluation made of what works and what doesn't work on a cost-benefit basis.

The quality issue, of course, has been exacerbated by the fact that women, who still constitute approximately two-thirds of the nation's classroom teachers, now have expanded career options. Talented and intellectually able women who in the past entered teaching because it was the traditional and logical thing for them to do are now entering other professional fields. Within the past decade or so, a "perverse form of indentured servitude" has been clearly weakened as many of our most talented women quite appropriately select a much broader range of career options. For example, in 1965 only 6.5% of medical school graduates were female, in 1980 the percentage soared to 23.4%; comparable figures for law school graduates were 3.2% in 1965 and 30.2% in 1980 respectively. In 1950, only an infinitesimal .3% of the engineering graduates were women, by 1980, 8.8% were female. It is reasonable to assume that many of these women doctors, lawyers and engineers in prior years would have been school teachers. Their loss to education, of course, and the continuing expansion of career opportunities for able women further compound the problem of attracting into teaching adequate numbers of our most intellectually able young people. *We should develop detailed information on this significant phenomenon which has such consequential ramifications for the quality of the teaching force in the years ahead*.

The Special and Acute Problems of Urban and Minority Youth

No litany of major human resource or educational issues can or should ignore some mention of the acute problems of urban and minority youth. The saliency of the demographic changes and concomitant growth in minority population were discussed at the outset of this paper. Special attention, however, must be focused upon the shocking reality that almost one-half of black teenagers have no jobs and that youth unemployment is higher than 20%, more than twice the overall unemployment figure for the country. Dropout rates are as high as 40% in many of our urban school districts. These numbers project starkly the danger of our urban schools becoming, in one college president's words, "warehouses for the angry or staging areas for anarchy". We need to analyze these issues even more carefully and collect information which will provide the basis for implementing badly needed remedies in our problem plagued urban districts.

In addition to the traditional morally compelling arguments for equity or the equalization of educational opportunity, there are now selfish, vested-interest reasons for wanting to improve big city schools whose student population will be 90% minority by 1990. As the demographic data reflect, our population is aging and shrinking cohorts of young people must be productive to generate the revenues necessary to support services in a changing economy. More of these youngsters obviously will be of minority origin and they represent valuable human resources which our nation and economy cannot afford to fritter away. Indeed, there will be a 20% reduction in the entering workforce in the 16 to 24 year old age bracket through 1990. These numbers indicate that we need to develop the capabilities of all of our youth to their fullest potential. We must substantially improve our 35 to 45% high school dropout rate and as one cynical realist says (hopefully with tongue in cheek), "let us forget justice and think profits and economic survival". Data on issues like school dropouts and the shrinking youth cohort should be gathered and disseminated more extensively so that there is more public consciousness of the collective stake the entire society has in ameliorating the urban crisis.

The Need to Define and Achieve Quality at All Educational Levels

Several other issues warrant at least brief mention. There is growing concern for quality in every facet of American life from cars to schools. The recent reports on education have decried, for example, "the rising tide of mediocrity" or the "lack of standards" in our schools. While acknowledging the undeniable need to bolster the academic quality of many of our programs, we cannot afford to retreat from our national concern with the issue of equity or access which has received such emphasis in the past few decades. Quality does not preclude equity or vice versa. These two transcendent goals are complementary and support each other. The consequences to our society of emphasizing one at the expense of the other are grave because of the economic and demographic imperatives discussed earlier which highlight the dramatic increases in young minority group members and the continuing commitment of equalizing educational opportunity which must be at the core of our democratic system. At the same time, however, the quality of many of our institutions has become shoddy and a parallel commitment to quality is essential if we are to compete economically and educationally. Data on both the quality and equity issues must be persistently and extensively gathered so that enlightened public policies can be formulated to maximize chances for achieving these vital interrelated goals.

The Weaknesses of Foreign Language Instruction

Another issue warranting some mention is the appalling decline in language training in the United States in a shrinking world with an interdependent global economy. A persuasive case can be made that our competitive position economically is weakened because of the deterioration in both the quality and quantity of our foreign language instruction. Many Japanese and Europeans, for example, know and speak our language. The numbers of our citizens able to converse in other tongues is abysmally limited. Our businesses thus have a distinct disadvantage in areas like marketing and sales vis-a-vis their foreign competitors. The statistics tell the story more graphically than words. Only 15% of our high school students currently study foreign languages, the figure was 24% in 1965. Only 5% of high school students study foreign languages beyond the second year and just 8% of our institutions of higher education require students to have studied a foreign language; in 1966, 34% of the institutions required such instruction for admission. These shocking numbers should encourage us to welcome many new citizens who have the advantage of speaking languages other than English. These new residents who can speak Spanish and other languages can prove to be a valuable economic resource as we seek to improve our international economic position. The continuous collection and dissemination of data which reflect this crisis in foreign language instruction hopefully will ultimately improve a serious situation.

Relationships Between Elementary-Secondary and Higher Education

Another issue requiring the attention of educational leaders and data collectors is the need to more closely coordinate elementary-secondary and higher education. The twelfth grade, of course, is the magical and mystical dividing line between the two and efforts must be expanded to more effectively bridge the levels in areas where they overlap. We can no longer afford duplicative programs where they may exist. For example, many community colleges offer remedial and vocational programs that are quite similar to the offerings of school systems under the aegis of adult basic education. Mechanisms should be created at the local or regional level that will coordinate programs more effectively. Also, we ought to move away from categorical set-asides and blatant political trade-offs that temporarily "buy" peace between the levels. If educators do not take the initiative in rationalizing their systems, politicians certainly will at a time of declining resources. Information about activities which bridge the levels such as the College Board's Project Equality should be compiled and widely disseminated.

There are major pragmatic reasons for increased communication between higher and elementary-secondary education. As the competition for dollars becomes keener, there is a real danger of dysfunctional interlevel conflict both for governmental and private funds. In California, for example, the recent comprehensive reform legislation included an $800 million tax increase to be used for the schools; higher education and other services were cut. Even in the area of private giving where higher higher education has had the terrain to itself, there have been developments that could generate negative competition between the levels. The public schools have just recently turned to the private sector for support and local and regional foundations are being created throughout the country to serve as the conduits for contributions. A new Public Education Fund has the explicit purpose of serving as a catalyst in the development of local and regional foundations which will serve as mechanisms to funnel private resources into the schools. Data on the dollars raised at the various levels should be collected and trends analyzed.

Although there has been little or no overt interlevel conflict thus far, the dangers of this occurring are all too real. Politicians are the first to say that internecine conflict within a policy realm like education could hurt everyone, particularly at a time when resources will be harder to find and competition is keener. Educational leaders from all segments of the enterprise have a responsibility to try to ensure that such interlevel conflict does not transpire.

The Appropriate Federal Role in the Shaping of Educational Policy

The appropriate role of the federal government in the shaping of educational policy remains a persistent and controversial issue. International political and economic issues and concerns would appear to require national responses if not huge federal programs. It seems unlikely that 50 state school systems, some 16,000 local school systems, and more than 3,200 post secondary educational institutions have either the human or economic resources for the expensive research and development investments that have to be made in the new technologies, science, math, foreign language and computer education. The necessary technological crash programs and heavy up-front investments needed in new areas like computer education must be promulgated at the federal level as we compete in an international economy against countries like Japan and West Germany which have national planning. Efforts should be made to consistently assess the capabilities of each level of government to provide new technologies in the instructional process.

Although there were without doubt excesses in some of the educational initiatives undertaken by the federal government in recent years, we cannot afford to throw the baby out with the bathwater. In other words, we need to find some middle ground between federal domination and a programmatic void in technical areas which the country cannot afford. Indeed, our concerns with areas like math, science and technology ought to serve as lightning rods as all segments of the society acknowledge a common stake in national priority setting and a balanced federalism if we are to maximize the development of our human capital.

The major issue, in essence, is how we reconfigure a fragmented and uncoordinated educational and training system to meet emerging economic and workforce needs without stifling the strength of its diversity. A key component in this reconfiguration, as I have stressed, is widespread acceptance of a more generic and broader definition of education as representing more than just traditional schools and colleges which serve the young. Such a reconfiguration will have to be based on much more mutual understanding and exchange of information between and among the varied educational delivery systems. This understanding and these information exchanges will require new data bases on a wide array of non-traditional or non-institutional education service providers.

In Conclusion

In this paper, I have attempted to provide an environmental scan of a number of salient social, political and economic issues which impinge upon educational policymaking. Indeed, my overarching thesis is that schools are the dependent variables of larger social and economic forces which drive public and educational policy. NCES ought to broaden and diversify its approach to the collection and dissemination of information so that data relating to these larger environmental factors which so inordinately influence schools are woven into and amplify straight statistics pertaining more narrowly to school systems and educational institutions.

More accurate and broadly gauged data will be essential if elementary and secondary education is to maintain political and economic support at a time when less than 30% of the adult population has children enrolled in school. The recent reform movement offers unique opportunities because of the interest manifested in improving education by Governors, business and civic leaders and other political influentials. Economic support will be forthcoming, however, only if taxpayers are convinced rather quickly that there are tangible results of their increased investment as measured by improved student performance.

In fact, the reform movement may already have crested and there is an urgent need within the immediate future to sort out the various reforms and determine which have been most effective on a cost-benefit basis. Michael Kirst of Stanford University identifies the following four phases in the issue attention cycle: (1) alarmed discovery; (2) crisis activity; (3) disillusionment with results; and (4) return to neglect. A compelling case can be made that the education reform movement currently is rapidly passing through phase 2 and that its ability to endure will be predicated upon its effectiveness as measured in phase 3 which is impending.

Thus, NCES' current effort to redesign its elementary and secondary education data program comes at a particularly propitious time. Its data collection strategies should be tied in whenever possible to the reform movement which is sweeping the nation. Systematic data should be compiled which will determine which programs are successful and which are unsuccessful because the resources will not be available to support the entire panoply of reforms that have been offered. It is important to document as expeditiously as possible some of the results which appear to be salutary. Kirst, for example, points out that in California there are indications that there already has been a 20% increase in high school science and math enrollments since the comprehensive reform legislation was recently enacted.

The following represents a very brief non-inclusive summary of a number of the major suggestions made in this paper:

1. A broader definition of education must be assumed to include other deliverers of education such as the private sector, the military, voluntary associations and many other providers of education and training services.
2. Greater communication and collaboration between and among traditional schools and other providers of educational services should be encouraged.
3. Data should be gathered from all deliverers of educational services not just traditional schools. Comparisons and exchanges of information should be made between the diverse sectors.
4. Data should consist of more than just straight school-focused statistics. Sources of information should include case studies, compendia of promising practices and processes, and examples of successful teaching and learning processes from all segments of the education and training world.
5. Data should be packaged in more diverse and creative ways rather than through monolithic statistical formats. Quick, incisive summaries should be provided and aggregate data should be supplemented by information from the building and district level within school systems.
6. In addition to the customary school related data, NCES reports should stress the collection and dissemination of more diversified and additional information in areas like
    - (a) demography (age, race, region)
    - (b) economic trends and employment opportunities
    - (c) dropout rates
    - (d) social data (single parents, latchkey kids)
    - (e) quality and quantity of Math and Science teachers
    - (f) success or failure of various differential staffing plans
    - (g) status of equity and quality issues and their relationship
    - (h) impact of technology
    - (i) status of foreign language instruction

HISPANICS AND EDUCATION DATA

Rafael Valdivieso

HISPANIC POLICY DEVELOPMENT PROJECT

Prepared for the National Center for Education Statistics, July 19, 1985

HISPANICS AND EDUCATION DATA

Since the late 1970's the National Center for Education Statistics (NCES) has produced some of the most useful and usable data about Hispanics from among a wide variety of governmental agencies. But that leadership edge is beginning to erode as other statistical agencies such as in the Department of Health and Human Services have come to realize their data gaps and are mounting serious data collection efforts about Hispanics. The momentum for these efforts were accelerated after the 1980 Census documented the growing number of Hispanics.

NCES's data efforts in regard to Hispanics have been in both illuminating specific problems such as language use and background as well as actively pursuing strategies to amplify the collection and dissemination of Hispanic data. Ironically, while we now know more about Hispanic high school students in general and about language use in specific, we still have not developed adequate data and models to explain how English language deficits or bilingualism interact with other factors to affect attainment and achievement. On the other hand, because of the NCES data, we can now identify a number of basic issues beyond the traditional concern of language.

But NCES cannot rest on its laurels and, in fact, will need to move aggressively to maintain its current assets and to anticipate new demands for data on Hispanics.

It will be another two decades before Hispanics actually outnumber blacks to become America's largest minority but, of course, in a number of cities and in some regions this is already a reality. More critical in regard to data needs is that Hispanics are replacing blacks as the group at the bottom of the education ladder in terms of both attainment and achievement. We know Hispanics drop out at greater numbers on the different levels of education but that Hispanics achieve about the same as blacks now is not as readily known. In comparisons of black and Hispanic achievement, the fact that larger numbers of poor achieving blacks are still in school, and therefore, lower their average scores is usually not considered.

EDUCATION LEVELS AND TRANSITIONS

Pre-School to Early Elementary

NCES will soon probably have to expand its data collection to the pre-school level (5,4 and 3 year-olds) in order to account for the effect on later schooling of the following (often, interrelated) societal factors:

(1) More women entering the work force while their children are young;

(2) The recent return of high levels of poverty among children, especially minority;

(3) Increase in female-headed households; and

(4) The push for early childhood education as a major policy response to these three factors.

HISPANICS AND EDUCATION DATA

What is the Hispanic situation for each of these factors and what are the implications of these factors for collecting data?

Although it is a recent change, Hispanic women (49%) are participating in the labor force at rates comparable to white women (49%) and black women (53%) (DeNavas and Fernandez, 1984).

Since 1979 the poverty rate among Hispanics has been climbing steadily. The poverty rate for Hispanics under 18 in 1983 was 38.2 percent, compared to 17.3 percent for Anglos (Pear, 1984). Furthermore, Hispanic children who live in female-headed families are more likely to live in poverty than either white or black children in the same situation. The respective figures are 70.5 percent, 47.6 percent, and 68.5 percent. More black children, however, live in female-headed households although even that dubious distinction has almost been matched by one Hispanic sub-group: Puerto Ricans.

In 1982 about 45 percent of all Puerto Rican families were maintained by a woman with no husband present, compared to an average for non-Hispanic families of 15 percent. Put another way, 55 percent of all Puerto Rican children live in single-parent households (Bould, 1985). About a third of the Puerto Rican families had one worker and other third had no workers. Given this set of circumstances, one can readily understand why 42 percent of all Puerto Rican families live in poverty.

Overall, in 1983, 23% of Hispanic families were maintained by women (U.S. Department of Commerce, 1984) and the median money income in 1982 for such women was $7,436 to $7,458 and $13,496 for black and white women, respectively. (DeNavas and Fernandez).

Probably, the single best background correlate to a child's education achievement has been the mother's education. The low education attainment of Hispanic women, especially those in poverty, complicates the pre-school situation for their children. Among Hispanic families headed by a female under 45 years of age with only children under 6 years, 64% had less than a high school education in 1984, or put another way, 28% had only an elementary school education. The comparable figures for less than high school education among white and black females are, respectively, 26% and 25%. (U.S. Department of Commerce, 1985a, Table 8).

What is it like to be raised in a poor family headed by one adult? In the past, poor children entered school at a disadvantage because of less-developed verbal competence than a child coming from a middle-class family. What further deficits will the children from one adult homes have if they have even less exposure to adult speech (as limited as it may be). In the case of Hispanics, we have a further complication in that Hispanic children often come to school speaking mainly Spanish or a combination of Spanish and English.

The larger policy question to address is whether we are going to have, as Bud Hodgkinson has said, more "damaged goods" on arrival at the doorsteps of our elementary schools that we have had in the recent past. Or, should we be trying other policy approaches to prevent or reduce this likelihood. Policy makers will not be able to decide without adequate data.

Because of these grim statistics and the knowledge that pre-school programs have had quite positive long-term effects with disadvantaged children, pre-school programs and kindergarten for these children should be widely adopted. Head Start and pre-kindergarten programs for children from low income families are seen as a way to get these children ready for school as well as to identify and correct a whole series of handicapping conditions early before they are compounded in elementary school and later life. Some of the handicapping conditions and the neglect of these conditions are directly related to poverty and poor health care for the mother and child.

Yet, Hispanics did not participate as heavily as other groups during the late 1960's and 1970s in federal programs such as Head Start. Even today, while 37.6 percent of white and 36.3 of black 3 and 4 year-olds are in preschool programs, only 23.5 percent of Hispanic 3 and 4 year-olds are in similar programs (U.S. Department of Commerce,1984). We do not know why this is the case, except that Hispanic families may prefer not to entrust the care of young children to persons who are not relatives. The Current Population Survey could ask why a child is not enrolled at the appropriate level of schooling. The question would yield useful information on dropouts at the other age levels. We will return to this question in the next section.

As for NCES collecting data at the pre-school level, I am not familiar enough with data collection methodologies at this level to offer advice, except that a longitudinal survey, with interviews of teachers and parents, to cover the transitions from pre-kindergarten to the early elementary school years should be considered.

Upper Elementary to Early High School

Excluding the elementary school years, Hispanics are less in school, including the pre-school level, than either black or white. The Hispanic attrition rate is higher than either at every level after elementary school. We know Hispanics are often retained a grade in elementary school. We also know that tracking becomes more prominent from the upper elementary school years on.

One might well ask at this point: how does the problem of language fit into this situation? Isn't the problem of limited proficiency in English the main cause of low academic achievement and dropping-out for Hispanics? A few comments on this topic are in order. Yes, langauge does play a part, but the situation is more complex than would seem apparent. By the time Hispanics reach the ninth or tenth grade, their language difficulties in earlier years may have caused them to be retained a year or two in earlier grades. Almost 25 percent of all Hispanics enter high school overage (Brown, Rosen, Hill and Olivas, 1980). Hence, they are behind their age contemporaries in school and ahead of their grade peers in physical and emotional development. Combined with other factors as poor grades and attraction to work, being overage frequently results in students dropping out of school.

Yet, because the complexity of this situation is not usually captured in surveys of drop-outs, the language factor does not loom as large in the survey results. More data and research is needed to tease out the factors and their interrelations in this situation.

As useful as the High School and Beyond data has been for studying dropping-out between the sophomore and senior years, it can't provide us with the data we need to study the interrelation mentioned because High School and Beyond does not contain, of course, any data on those students who dropped out before the spring semester of the tenth grade. As mentioned above, many Hispanics are behind their age peers in school and reach age 16 before they get to the tenth grade. According to a report prepared for NCES by Hirano-Nakanishi (1983), about 40 percent of all Hispanic students who leave school would do so before reaching their sophomore year.

Just as High School and Beyond spans the important transition from high school to post-secondary education and the world of work, we need a longitudinal survey that spans from the upper elementary through junior high school years to high school, i.e., 4th, 6th, 8th and 10th grades.

Adult Education

Forty percent of Hispanic 20 to 24 year-olds are not high school graduates and the comparable Puerto Rican figure is 46 percent (U.S. Department of Commerce, 1985). We can also assume that some of these young adults are already married and have children, and for many others, this will soon be the case. It is important to know what these individuals plan to do about completing high school or obtaining a G.E.D.

What complicates the out-of-school/non-graduate statistic for young Hispanic adults is that it includes young Hispanics who have recently immigrated into this country and are not high school graduates. We are not at all sure about an estimate for this group. In any case, the larger the number of young Hispanic adults who have recently arrived in this county without a high school diploma, the more it is an issue for public policy.

While it would be difficult for NCES to directly address these issues through its own data collection efforts, the October Current Population Survey is a good alternative. As mentioned in the pre-school section, I think a question or two could be included that would ask: If you or children in your household are not enrolled in school at the appropriate level, why are you and/or they not?, and do you/they plan to continue your education.

Even the growing phenomenon of individuals "stopping-out" (or temporarily stopping their education with intentions to return) at the post-secondary level, and increasingly at the secondary level, can be addressed through these questions. These questions would yield valuable information for higher education planning.

Another alternative for collecting these data would be the adult sample of NAEP.

HISPANICS AND EDUCATION DATA

SPECIFIC DATA NEEDS OR MODIFICATIONS

The following comments on classificatory variables and levels, especially those more pertinent to Hispanic than other groups, are offered as short advisories that also contain suggestions for modifications or continued use of the variables or levels.

Ethnic Sub-Groups. While Hispanics of the various groups are beginning to coalesce on some issues, the socioeconomic and education characteristics of the different groups are as diverse as ever and are beginning to diverge even more so from each other. The residual category of Other Hispanics is now the second largest group of Hispanics and has become a hodgepodge of the various, most recent Hispanics to the country with the oldest Hispanics in the country, the Hispanos in the southwest. I mentioned this only as an alert as I think the Census Bureau will need to do something about this category before other agencies can make any changes.

Nativity and recency of migration. While a large majority of Hispanic students, and even their parents, are born and raised in this country, continuing immigration still makes these important variables to maintain. In fact, if political violence continues in Central America, and possibly escalate in Mexico because of its economic and political dislocations, then we will see even greater waves of adolescents and young adult immigrants to this country. Right now we are experiencing a surge of new Salvadoran adolescent male immigrants because the draft age in El Salvador has been lowered to 12. Many of these youngsters have not had any schooling in their country.

Language. Since the mid-seventies we have developed ways to identify language minority students and have learned more about their problems in school. We have moved towards core definitions of both language minority target groups and the services they receive.

While we have a considerable amount of self-reported data from national data bases as the 1980 Census and the High School and Beyond, etc., a short assessment of reading proficiency in both English and, possibly, Spanish, would give us a better grasp of the salient features in the self-reported assessments. In general, the combination of these and other activities will allow for cross-walking" and comparative analyses with other data bases. In this sense, the proposed assessments would further enhance the federal investment that has already been made in this area. The proposed assessments are the next stage in a logical progression or expansion of inquiry in to the education of language minority students.

I understand ETS proposed to add these assessments and an expansion of the language minority construct to the NAEP data base, but I do not know the outcome of this proposal. Such data and analyses would move us closer to understanding the interrelationships between language difficulties, school policies and practices, school age delay and dropping-out. Because of the extraordinary high drop-out rates among language-minority Hispanics, research in this area is a critical research priority.

Gender. While white and black gender differences are moderating, Hispanic gender differences are still quite sharp.

Age. Age has already been mentioned in regard to its importance in studying Hispanics who are over-age for their grade placements.

Parent education. NCES used "Less than High School" as its first category on parent education in HS&B and other surveys. This is not adequate for Hispanics and their low education attainment. The first category reported should be "6th grade or less," or possibly, "8th grade or less." For Hispanics, 25 to 34 years old, 26.4 percent had completed less than 9th grade in 1982 (U.S. Department of Education, 1984).

"At-risk" construct. NCES should consider developing a construct for at-risk students, using below-average grades and achievement scores in the first quartile and perhaps some other conditions.

The construct would be helpful for addressing a variety of policy concerns which will arise as we enter the second stage of the current reform movement.

NCES Data Sets to Keep or Delete

From my perspective, High School and Beyond is the most important to maintain because of its indisputable value in assessing student performance, school effectiveness and student changes over time.

I can see the value of the other NCES data bases but I am not nearly as familiar with them as I am with HS&B.

The Common Core of Data is an important collection as well as concept. In spite of the great difficulties involved, more should be done to obtain the use of common definitions among the States for dropping out and other student outcomes. Likewise, more effort should be given to collecting data by racial/Hispanic indicators even though substantial political problems may be involved. This same comment applies to most NCES census and sample surveys, including the needed private school survey. About 10% of the Hispanic student population is in private schools.

In view of similar information that could be collected for items like vocational education facilities and equipment or scientific equipment and materials, why is information on library/media centers collected? Is this survey necessary?

Other Comments and Recommendations

Census Data by School Districts

To my knowledge, not much use, at least by most people outside the government is being made of census data organized by school district. Perhaps this is already the case, but if it is not, an NCES program officer should be assigned the task of working with the Census Bureau to make data organized by school districts better known and accessible to policy analysts and researchers in education. The program officer should also be involved in assembling as well as devising solutions to technical problems in using these data.

Multiple Sources of Survey Data

A major strength of the HS&B study is the confluence or triangulation of data obtained through different methods or sources. This approach should be praised and continued in future surveys. In this regard, an expanded parent or home interview for HS&B would have been useful and should definitely be pursued in future surveys.

Oversampling

The oversampling of some Hispanic subgroups in HS&B was critical to making this survey the most useful national survey ever. Again, this approach should be continued in future surveys.

Dissemination of Data About Hispanics

NCES has not kept pace with its previous achievements in this area. The publication of the Condition of Hispanic Education by NCES in 1980 was widely recognized and applauded in the Hispanic community. In view of the new data collected and the pressing needs for this information, an updated edition of this publication should be on NCES's agenda for the coming year.

NCES should also consider preparing and releasing for the use of analysts and researchers an index of NCES reports and bulletins, contractors reports, and tabulations that include Hispanic indicators or categories.

In regard to developing and presenting tabulations by race and ethnicity, the collapsing of Hispanics with other groups should be avoided if at all possible. For example, as insightful as a recent NCES report, An Analysis of Course-Taking Patterns in Secondary Schools as Related to Student Characteristics, was for the study of course-taking, it was marred, because it collapsed Hispanic with Others.

The State Education Statistics wall chart is quite handy, especially the challenge goals, but it should include percentages of single parent households and limited-English-proficient children under population characteristics.

REFERENCES

Bould, Sally
Unpublished tabulation from March Current Population Survey of Household and Family Characteristics.

Brown, George H.; Rosen, Nan L.; Hill, Susan T.; and Olivas, Michael A.
(1980) The Condition of Education for Hispanic Americans. National Center for Education Statistics. Washington, D.C.

DeNavas, Carmen and Fernandez, Edward.
(7/27/84) Condition of Hispanic Women in America Today. (A presentation) U.S. Department of Commerce, Bureau of the Census).

Hirano-Nakanishi, Marsha.
(1983) Hispanic School Dropouts: the Extent and Relevance or Pre-High School Attrition and Delayed Education. (Report No. 17). Los Angeles: National Center for Bilingual Research.

Hodgkinson, Harold L.
(1985) All One System: Demegraphics of Education, Kindergarten Through Graduate School. Washington, D.C.: Institute for Educational Leadership, Inc.

Pear, Robert
(1984) "Rate of Poverty Found to Persist in Face of Gains." New York Times. August 3.

U.S. Department of Commerce, Bureau of the Census·
(1985a) Household and Family Characteristics: March 1984. Current Population Reports, Series P-20, No. 398. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Commerce, Bureau of the Census
(1985) Persons of Spanish Origin in the United States: March 1982. Current Population Reports, Series P-20, No. 396. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Commerce, Bureau of the Census
(1984) Social Characteristics of Students: October 1983. Current Population Reports, Series P-20, No. 344. Washington, D.C.: U.S. Government Printing Office.

U.S. Department of Education, National Center for Education Statistics
(1984) The Condition of Education. Washington, D.C.: U.S. Government Printing Office.

National Statistics to Improve Educational Productivity

Herbert J. Walberg

University of Illinois at Chicago

Now is the time for the U.S. Department of Education to expand and improve national, state, and local statistics collected on education. With justification, many national reports, including A Nation at Risk (U. S. Commission on Excellence in Education, 1983), urge educational reforms; and state legislators and other groups are beginning substantial changes in educational policies. Not only is it in our national interest to study the effectiveness and efficiency of education, but the magnitude and effects of current reforms deserve the closest scrutiny during the next decade.

The government role in education may be more complicated in the U.S. than in other countries; the unique qualities of our system may both help and hinder the effectiveness of education and the collection and use of statistics. Unlike most other countries, we have no centralized governing and data-gathering ministery of education; (education policy is largely a state and local responsibility). Nor do we have, like many other advanced countries, a single centralized national agency to collect and analyze statistics on education, business, health, agricultural, cultural, and other matters.

It is not my purpose here to take sides on complicated questions of federal, state, and local control of education; nor, for that matter, on the proper influence of school administrators, teachers, parents, and students on educational goals and means. It should be recognized, however, that the the U.S. Office of Education (predecessor of the Department of Education) was created by Congress earlier in the century to collect statistics with a view toward improving education. Today, moreover, special interest and political groups ranging from liberal to conservative agree that valid data are required to formulate effective educational policies (Cooke, Ginsburg, and Smith, 1985).

Meter Sticks for Education

The value of of statistical research depends on valid comparisons, several of which can usefully influence educational policy. These include comparisons of the U.S. with other countries; among the 50 states; and among students, classes, schools, districts, and regions. In principle, all of these may be "cross-sectionally" compared at a single point in time, or changes in them may be "longitudinally" compared across years.

Even though education spending is one of the larger and growing fractions of the gross national product, and even though investments in "human capital" may be vital to future welfare (Walberg, 1983, 1984), education statistics are fundamentally invalid in several ways. The most important flaw is the lack of universalistic, absolute measures of the primary goal of education--learning.

Since Alfred Binet's turn-of-the century precedent, test developers have normatively compared students to one another rather than to absolute standards of performance. They have developed tests to reveal differences within homogeneous groups of students by selecting items that about half the students can correctly answer and which yield "normal," bell-shaped distributions (or positively-skewed distributions from easier "mastery" test intended to detect slower learners for remediation). Whether teacher-made or commercially-developed, such tests may be administered to intact classes or even larger groups to measure learning fairly accurately but over a narrow range.

A typical range of correct answers on a 40-item test might be from 20 to 40 points correct, or a ratio of two to one between highest and lowest scores. But some students might have guessed the correct answers on a fourth of the questions and deserved lower scores; and others might have been able to demonstrate more impressive knowledge had the test been longer or composed of more difficult items suited to their ability.

Such tests are reasonable if the purpose to compare students within a narrow group with one another. Researchers might, for example, compare students within a grade level, or a school district or state on standardized, commercial tests. Similarly, teachers ordinarily compare students on material covered in their lessons: Their students may be ranked relative to each other; or a percentage correct score may be calculated. Such practices are ingrained among educators and test developers; and it may be be difficult to think that learning and other accomplishments could be distributed other than a normal curve in a narrow range. A little reflection, however, about the following question and tasks illustrates the contrary.

What distribution will result from the following tasks? Name the streets of Chicago. Throw a discus as far as you can. Prove mathematical theorems. Name the presidents of the U.S. in order. Name the capitals of Asian nations. Write Urdu verbs and French nouns. Play as many tunes as you can on the violin. Give the rules of National Collegiate Debate. Show how much weight you can bench press. Demonstrate chess openings. Put your pocket money on the table. Run as far as you can at the rate of four minutes per mile. Count the number of times you have been in trouble with the police. Enumerate the articles you have written in the school newspaper. List the countries you have visited. Repeat the numbers read to you.

Obviously, few students could approach maximum human performance on these tasks. Indeed, many would attain a score of zero on most tasks, although a few would far exceed the mean perhaps by a multiple of five, ten, or more. Human performance varies considerably more than the usual norm-referenced tests can reveal.

Absolute Measures

These distributions are of measures or counts of instances of what might be called absolute measures on ratio scales. They have a definite zero point and can range over magnitudes, some indefinitely. They are similar to the absolute measures such as meters and kilograms in the natural sciences, time and money in economics, counts and measures of behavior and perception in experimental psychology, and scales of ordinary experience. Educational research, in comparing individuals to means, percentiles, and other relativistic norms of groups, has denied itself absolute measures that have made for fundamental understandings and comparable empirical results of the natural sciences, and possibly great increases in productivity such as those in agriculture and industry.

The eminent psychometrist John B. Carroll (1982) finds little fundamental progress in mental testing since Thurstone in 1925 (pp. 67, 77). Had psychometrists continued Thurstone's efforts to calibrate mental abilities and tasks to chronological age and learning time, however, educational productivity might be much better understood and optimized today, a half century later. Before research can establish the magnitude of the dependence of academic learning on its causal factors rather than simply the sign of the dependence, it may be necessary to develop adaptive, cardinal measures of learning and its factors over many magnitudes. Such a development, following productivity research in agriculture, health and industry, might lead to similarly great strides in enlarging human achievement and accomplishments.

Oddly, the models that might be followed to develop absolute educational measures are outside the mainstream of academic measurement. Athletic coaches and fans undoubtedly have the most accurate and wide-ranging statistics; they can compare speeds, weights, ratios of wins and losses, and other absolute indicators. They can, for example, easily compare the times of a 100-meter dash of people of different ages, countries, and years. So also can someone who has devoted little time to studying sprints times; comparisons and their meanings can be readily understood by non-specialists.

New training regimens in athletics can be evaluated and individual performance over time can be assessed according to absolute measures. It seems that the century's steady progress in athletics--for example, the Olympics--is in part attributable to well-defined, absolute scales accurate to many significant digits across wide variations in performance. Typing tests of speed corrected for accuracy offer another useful precedent.

What about academic performance? What is lacking are absolute standards or measures that would enable us to compare children of different ages, grade levels, and abilities. It is though each test publisher and teacher had a different meter stick; and yet there is no way to equate them. Thus, test scores in a California district can not be validly compared with a New York district unless the same test happened to be administered in both districts. If a single district switched tests, it would be unable to compare earlier and subsequent scores.

If the test publisher changed editions or "re-normed" the test, then comparisons of earlier and subsequent scores are likely to be invalid (even the Scholastic Aptitude Test appears to have drifted in difficulty over the last few decades). Similarly, we cannot compare the performance of third and sixth grade students to find out how much they have learned because comparisons are strictly valid for only students who have taken the same test (not the forms for earlier and later grade levels); and the tests are capable of measuring only a narrow range of a few grade level equivalents.

One possible solution is to calibrate items and tests to national standard tests. The National Assessment of Educational Progress may offer a reasonable basis of national standards; but it represents only three age-levels, 9-, 13-, and 17-year olds. Expanding the National Assessment and coordinating it with other large-scale testing programs could lead to a more accurate picture of U.S. achievement and the possibility of a universal "meter-sticks" of learning to which other measurements may be calibrated. Given calibration formulas and procedures, states and school district staff, citizens, parents, and students could compare their scores with their previous performance or with the progress of others.

Computer-Assisted and Adaptive Testing

Even NAEP-based calibration, however, would adhere to this century's convention of giving each child within a class or grade the same test, under what would be called "batch processing" in industry. A far more efficient and time-saving approach is "tailored-testing" (see Carroll, 1982) which flexibly adapts test items to students over great ranges of ability (rather than the reverse).

For several decades, it has been possible both in principle and in practice to program computers to assign the most discriminating items to each student, based upon her or his prior responses during the testing session (Carroll, 1982). As few as 15 tailored items can yield scores as reliable as 90 batched items suited to the average student. Alternatively, 90 tailored items given in two hours can yield very accurate assessments not in one subject but in all the major subjects of the standard curriculum. Or, 90 items could provide highly detailed assessments of skills in a single discipline, for example, word choice, grammar, spelling, and punctuation in written composition.

The increased efficiency in time use and the computer's capacity to record large amounts of information make it feasible to monitor individual student progress more frequently, accurately, and comprehensively. With a thorough, continuing assessment of what each student need to learn, it should become equally feasible to provide computer-adapted or tailored instruction. Such instruction is by no means a panacea, but it is among those educational methods that provide moderately superior achievement; and it has the further advantage of saving students' study time (Walberg, 1984). It can be expected that hardware costs will continue to fall, while software increases in sophistication and interest.

There is no reason why schools alone should provide computer-based assessment and instruction. It might be argued that since the schools have changed their basic technology of explanation, recitation, and seatwork very little since the turn of the century, other agencies might also be given an opportunity to explore these new opportunities.

In principle, students and parents could monitor student progress on absolute scales provided by entrepreneurial public and competing private groups. At a current, one-time cost of a few hundred dollars for a "dumb terminal" (without programs and memory) and a modem (to convert telephone-acoustic and computer-electronic signals), they could call a large, "mainframe" computer from their homes an 800 or a local number, take a tailored test on any subject, and bill the cost of a few dollars per assessment to a private credit card.

Public and private schools, state departments of education, and proprietary corporations could provide not only assessments but instruction as well in this way. A state and local community could finance such a system by providing an education credit card worth, say, $3,000 per year to be spent on educational services parents and students chose. These might include a mix of home and school computer instruction as well as conventional school instruction. State and local educators could suggest minimum competencies and hours of study, require performance levels for passing from one grade to the next and graduation, or impose a great number of regulations and certification practices.

Conventional and new services might range from traditional instruction in neighborhood public schools to computer-based education offered by in-state and out-of-state public and private schools and for-profit corporations. The services might be provided in schools, shopping centers, mom-and-pop neighborhood outlets, or in homes at any hour of the day or time of the year as the need or interest arose.

It is possible to program computers to monitor student progress in relation to activities on terminals and in other educational experiences. Automated statistical analyses can show which activities lead to the highest rates of learning. Nationally-calibrated absolute measures would enable states, local districts, teachers, parents, and students to weigh the costs and benefits of various activities in making their educational plans and choices.

Some may find such new forms of instruction very much to their liking; others may find them undesirable; that is human nature. Adaptive testing on absolute scales, however, can be made carried out separately from the means of instruction. It can provide convenient measurements to assess students and programs when they seem needed.

National Statistics by Computer

Wassily Leontief, Nobel-laureate economist and inventor of national input-analysis finds it most difficult to apply his methods here in the United States where he first developed them: "The United States is the only advanced country in the world that does not have a central statistical office. Each department of the government collects statistics in the area of its own particular interests. Users of such data spend much of their time trying to reconcile and align information coming from these different sources" (Leontief, Duchin, and Szyld, 1985, p. 419).

Yet, even within the Department of Education, statistics are uncalibrated, unsystematically collected and archived, and poorly analyzed to guide national and local educational policies. Another country that keeps close track of national progress offers an interesting example of what can be done.

Growing faster economically than Japan is Singapore, where physical resources are scarce and "human capital" is taken very seriously by Prime Minister Lee Kuan Yew and others in the central government. Officials in the national ministery of education can call up in an instant any student's or young adult's test records or a mass of them for comparison from national computer bases. They can find the most qualified person to fill a particular job, or tell an American entrepreneur or manufacturer the number of trained people available for various high-value-added, growth industries such as tourism, electronics, and petroleum processing.

It may be argued that the U.S. may be too big, diverse, complicated, and perhaps fixed in its educational ways to enact such an innovation. Besides, possible abridgment of freedom and confidentiality are likely to worry and deter educators from something so intrusive. Still, the Internal Revenue Service does not seem to have abused its vast powers to reveal confidential information; if anything, there seems more need to worry about the accuracy of the information supplied by taxpayers. There seems no good reason to think that measurement calibration and related services provided by the government would violate privacy.

Citizens, moreover, particularly poor and minority-group parents, seem more enthusiastic than the educational establishment for better measurements and higher standards. Conducted for business, civic, and minority organizations, for example, a recent poll of 1,816 Chicago residents showed 88 percent feel that all high school students in the U.S. should be be required to pass a standard examination before being graduated.

In addition, Chicagoans overwhelmingly favored a tougher curriculum for high schools: The lowest-income groups most favored extending the curriculum to more subjects; and blacks more often than others preferred the toughest requirements in science, history, and foreign languages for college- and non-college-bound students. Of Chicago adults with children in the public schools, 69 percent said they would send their children to private schools if they could afford it primarily because they would get a better teaching, attention, and discipline (Walberg and Hess, 1985).

If we start from the premises that we must inform citizens about their schools; that educators should be informed about their business including their costs, benefits, and views of citizens; and that better education statistics may help us to understand and solve our educational problems--then we need to think about harnessing the vast powers of the computer, as other industries have done, to increase competitiveness and productivity. School districts and state departments of education are enlisting computers in central offices and classrooms; and it would seem the proper role of the federal government to lend research support and technical assistance to help coordinate the efforts. Comparisons of all sorts will be more valid, other things being equal, to the extent that data are obtained uniformly.

In addition, national hook-ups, perhaps sponsored by the federal government, would make it feasible to conduct sample surveys of districts, schools, and students directly by computers. Students, for example, could rapidly complete tailored tests and questionnaires by terminal and modem. In compensation, they might be offered a small stipend, or at least they and their schools could receive an immediate summary of results which could also be provided by long-distance telephone connections to state and federal computers.

The further advantage is the speed at which such surveys and tests can be completed. The time-consuming steps of printing tests and questionnaires, mailing, key boarding and screening data, and the like can be skipped. Even analyses can be automated.

Like Gallop and other polls of 1,500 respondents that provide reasonably accurate estimates of public opinion in the nation, direct sampling by computer would make educational polls and national assessments fast and cheap; they would minimize the total human time answering questions yet provide more accurate estimates than far larger but unscientific surveys. Quarterly or even monthly survey reports on important measures could be made routine as they are in commerce and industry. Local, state or national assessments of special topics might be commissioned and completed in less than a month. In principle, we would not have to wait a year for the Phi Delta Kappan's Gallop Poll on education, several years for cycles of the National Assessment nor as much as a decade between International Studies.

A National Bureau of Educational Standards

Before turning to the kinds of measures that seem desirable to collect, it should be acknowledged that what is called for above is a tall order as compared with what has been planned and spent on educational research. What is spent on educational research by other standards, however, is minuscule. Even if spending on educational research amounted to $150 million annually (Walberg, 1983), it would be less than .006 percent of annual educational spending on public elementary and secondary education in recent years. By comparison, it is by no means unusual for growing corporations in competitive industries to spend 5 or 10 percent of annual revenues on research and development.

The costs of federal research on defense, space, and medicine obviously dwarf expenditures on educational research which may pay greater dividends for the nation's future welfare. As Adam Smith said and Japan demonstrates, human capital is just as important as physical and financial capital in determining the wealth of nations. And it is clear that education can be made much more productive in increasing the ratio of its benefits to costs.

To sustain the coherent programmatic data collection that seems required may require a agency of the U. S. Department of Education. Such an agency, perhaps called the "National Bureau of Educational Standards," would be analogous to that first created in Paris long ago for keeping standards of weights and lengths. Such an agency would have to be carefully planned and under close scientific scrutiny since it would have to provide precise definitions and measurements of education and learning, which tend toward imprecision, non-comparability, and intractability.

Such an agency would need to avoid partisan stances, value judgments, and declarations of what constitutes adequacy or excellence. Like the National Bureau of Standards which provides physical standards for our country, it would have to adhere to scientific and factual questions rather than values stances, inasmuch as it is possible in education.

It is the charge of the Department of Education to collect statistics. It appears, moreover, that no other agency, public or private, could take on the large task of thinking through, commissioning, and monitoring or conducting the research required to put such an agency in place.

Aside from calibration, a National Bureau of Educational Standards could serve as the central government repository and publisher of statistics on education in the U.S and, where appropriate and feasible, in other countries. In addition, for those who wish to analyze the raw data rather than examine pre-digested summaries, a National Bureau could serve, like the Library of Congress, as an archive of computer tapes of educational data that could be reproduced at cost by requests in writing, in person, or by telephone (including telephone requests for data transfers by computer).

It would have a capable technical staff to archive data in standard formats that could reproduced for secondary analysis by investigators in universities, state departments of education, schools, and newsrooms. In this sense, it could be modeled after the Institutional Consortium for Political and Social Research at the University of Michigan that archives and makes available major social surveys and public opinion polls. Although such surveys may costs hundreds of thousands if not millions of dollars, the tapes and codebooks for any can be supplied at cost a few hundred dollars. With artificial intelligence, it should become possible within a decade or two for non-technical people to query such data bases for a few dollars by voice over the telephone without having mastered programming.

The beginnings of these functions are already represented in the National Center for Educational Statistics that distributes the High School and Beyond data on about 58,000 sophomores and seniors, their parents, and teachers. In addition, the Educational Commission of the States made available and Educational Testing Service currently makes available at cost data on several-hundred thousand 9-, 13-, 17-year-olds and young adults collected in the National Assessment of Educational Progress. In addition, the tapes from the many surveys of nearly 50 countries participating in the International Association for the Evaluation of Educational Achievement could easily be copied and archived, as could other surveys such as the Gallop Polls on education, the Equality of Educational Opportunity survey, the National Longitudinal Survey, and the General Social Survey. These data sets were assembled during at a cost of perhaps $500 million and are largely under-analyzed. A National Bureau could serve as archive, calibrator, synthesizer, reporter, and at-cost distributor or raw data and results.

Current Federal Statistics

It has been said that democracy is the worst government except for all other forms. The same may be true of current statistics the U.S. Department provides. Therefore, we should be loath to stop collecting any data series in education that has already been started, even on things that have seemed of little bearing on learning. If anything, we must expand the collection and coordination of data, and encourage scholars and others to analyze the costly and valuable statistics that are currently available.

Federal government spending on education statistics, however, is small by several standards. In school year 1982-83, for example, spending on public elementary and secondary schools in the U.S. by federal, state, and local government was respectively $56, 52, and 8 billion, which comes to a total of $116 billion, which is 4.5 percent of the $2.6 national income (Indicators, 1985, p. 22). If the federal government spent $100 million on better educational statistics, it would be amount to less than one-tenth of one percent of total educational spending on public schools and might increase efficiency by many billion. Given U.S. government spending of $1.4 billion on statistics (Alonso and Starr, 1985, 123), education's 4.5 percent share (based on the public school percent of national income) would be $63 million, in contrast to $8.7 million in current spending by the National Center for Educational Statistics. Higher spending should yield better statistics and make the "education industry" more comparable to agriculture, medicine, and various industries that base practice upon productivity comparisons.

Even the aggregate and crude numbers on costs and enrollments now compiled by the federal government from data supplied by the states can raise pointed questions. For example, the 1985 Indicators published by the U.S. Department of Education shows that an average of $2,948 was spent on each of the 39.6 million children in public schools in 1982-83 (p. 22). (By comparison, according to Feistritzer, 1985, the average per-student costs of Catholic schools in 1982-83 was $782; and private school tuition was $1,029.)

Public school teachers reported an average of 24 students in their classes; and the average ratio of students to all full-time equivalent teaching staff (which includes special teachers) was 19. So, depending on the estimate, per-student annual spending was $55 or $70 thousand per teacher (p. 30). Teacher earnings, however, were about $19 thousand (Indicators, 1985, p. 30), or only about a quarter or a third of total costs (actually the true fraction must be even smaller, since complete earnings which includes moonlighting and summer work are included as earnings).

Thus, indirect costs in public schools are apparently two or three times as high as the direct educational services provided by teachers. Where is this extraordinary amount of money going if not to teachers? Does it account for the approximate 500 percent increase in inflation-adjusted, per student costs since 1930? Has academic achievement gone up accordingly? Is what is being provided by indirect costs as valuable to students as teaching services? Does it help teachers to do their jobs more efficiently? Can it be going to physical facilities during a period of declining enrollments? If it is going to administrators, can such heavy bureaucratic spending be justified? Are federal and state governments creating local bureaucracies to deal with special programs and complex regulations? Do any of these explanations fit with corporate trends toward lean, competitive organizations with the most senior administrators close to customers rather than layered away from them by corpulent staffs?

Whatever the answers to these questions, the numbers themselves are provocative; they stimulate discussion and research. Such data should be easily accessible so that the public and educators can deliberate about them. It is important to keep accurate tabulations of them over the years so that we can better understand how the levels of learning are changing, what is and is not changing them, and what might be done to increase effectiveness. We should be reluctant to omit any just as the Library of Congress avoids dropping subscriptions of unfashionable magazines.

The Quality of Federal Statistics

Even though we should continue and expand the collection of statistical series, we must be duly cautious and consider, as Aristotle advised, the source. The British statistician Sir Josiah Charles Stamp (1880-1941) warned:

> "The government are very keen on amassing statistics. They collect them, raise them to the nth power, take the cube root and prepare wonderful diagrams. But you must never forget that every one of these figures comes in the first instance from the village watchman, who just puts down what he damm pleases" (quoted in Alonso and Starr, 1985, p. 123).

Cooke, Ginsburg, and Smith (1985) compiled several alarming discrepancies in estimates in important national statistics on and related to education. School safety and security, for example, have found to be related to learning gains; but how safe are students? The National Crime Survey administered by the U. S. Department of Justice seems to indicate from household interviews that about 10 percent of junior and senior high school students are victims of assault, robbery, or theft each year. The National Institute of Education, however, reported from confidential answers by students in their classrooms that 10 percent were victimized each school month--an estimate at least ten times larger than the Department of Justice estimate based on parent interviews about their children.

Similarly, the U.S Department of Education's Vocational Education Data Systems reported 741 thousand New Jersey students taking high school vocational education courses in 1979--a number that exceeded the State's high school enrollment by more than 50 percent. In Virginia, the 29 thousand Indians which VEDS indicates as enrolled in vocational education represents more than three times the total Indian population, according to the State Indian Commission.

Even well researched variables in the mainstream of educational reform movement are suspect. Instructional time in the United States is rarely more than 60 percent of the school day; but the share varies by more than 2 to 1 among schools, and engaged time is only a varying fraction of allocated time, according to Cooke, Ginsburg, and Smith (1985). Japan's high schools may employ twice as much engaged time in the four years of high school (including extramural study) to yield achievement equivalent to the U. S. bacchalaureate in mathematics, science, geography, native and foreign languages, and music as well as non-academic pursuits (Walberg, Paschal, and Weinstein, 1985). It is, however, difficult to get more than ball park estimates of these important comparisons.

In Illinois, perhaps because of inflation of course titles or blurring of content, 80 percent of the high school students reported taking geometry, but a census of actual transcripts in the State revealed that only a quarter had. In California, 99 percent school attendance is reported; but students who have "valid" excuses there are reported as in school--a far different definition is given in other states (Cooke, Ginsburg and Smith, 1985).

It may be hoped that different means of reporting, biases of the reporters, and random factors may balance out and permit at least rough comparisons across respondents, states, nations, and time periods. But it remains a vague and often patently false hope; and the discrepant estimates of Cooke, Ginsburg, and Smith (1985) may chime the thirteenth hour on the educational statistics clock for some important figures. "Lies, damn lies, and statistics" said Benjamin Disraeli and Mark Twain.

The National Bureau Revisited

Science offers several ways to assess and solve such statistical problems; they deserve the support of a National Bureau of Educational Standards. One is to insist that highly detailed, explicit, and publicly accessible descriptions of data definition and collection procedures. Another is to commission papers and convene conferences to criticize, design, and redesign large national and international sample surveys. To some extent, these are major functions of refereed journals and associations in the natural sciences and their applications. They are often funded by U. S. government agencies such as the Department of Agriculture, Department of Health and Human Services, and the National Science Foundation in the such cases of expensive health and productivity surveys and massive projects in physics.

National groups of blue-ribbon layman or practicing professionals have tried to carry out these functions in preparing recent national reports on education. They have employed school visits, hearings, and a selection of expert testimony and papers. Democratic societies should allow, indeed, encourage non-technical deliberation and formulation of policy. Public and private commissions should deliberate and recommend goals, values, and means, which may be enacted by legislators, private agencies, and individuals.

But such groups should have accurate statistics as one basis of their deliberations. They may not have the technical competence to gather and assess the statistical facts; and technical experts may not be able to see beyond the facts to the public interest. In view of the limits of human time, skills, and knowledge, some specialization of function is required. A National Bureau of Educational Standards should be restricted to collection and assessment of data, calibrating and correlating measures, commissioning large-scale studies, making information available, and criticizing it. In this way, it may provide good data for policy analysts and decision makers.

It should, however, avoid political stances and recommending of policies and practices. The National Bureau of Standards accurately tells us how long yards and meters are, not how long our houses or apartments should be. The Department of Labor gives the incomes of occupational groups (with a margin of error) not evaluations of what income distributions should be; when it gets beyond the ascertainable facts to such predictions as the number of mechanical engineers required in ten years, it is often wrong. The Department of Agriculture can give the average corn yield of Iowa farmland and the increments associated with degrees of tillage, irrigation, and fertilizer; but the farmers decide how to farm. These seem instructive precedents.

What Educational Data Needs Collecting?

Following the lead of early agricultural experimentation, much educational research focuses on the relation of single causes and effects. Education, however, obviously involves many means and ends, each with an explicit or implicit cost or value. The promotion of efficiency requires the specification and measurement of the chief causes, means, or "factors" of production.

Experiments and statistical studies of productivity data together with cost and value estimates have enabled a wide variety of industries to increase the value of their output while simultaneously reducing costs thereby raising human welfare. Although such thinking may seem alien to some educators, the public ranks research on educational effectiveness higher in priority than most other fields of scientific investigation in the natural and social sciences (Walberg, 1983); and educators may do well to think more explicitly and unsentimentally about our business and to try to found it on the emerging consensus of scientific evidence.

It should also be said, however, that we educators are far from estimating explicit costs and values. The prior problem, now being solved, is estimating the magnitudes of effects of educational inputs on outputs, which primarily involves causal rather than value questions. It is these chief causes and effects that deserve first priority in national data archives.

Nine factors require optimization to increase affective, behavioral, and cognitive learning (see Walberg, 1984, and the cited references for a more detailed discussion). Potent, consistent, and widely generalizable, these nine factors fall into three groups:

Student aptitude includes:

1) Ability or prior achievement as measured by the usual standardized tests,
2) Development as indexed by chronological age or stage of maturation, and
3) Motivation or self concept as indicated by personality tests or the student's willingness to persevere intensively on learning tasks.

Instruction includes:

4) the amount of time students engage in learning and
5) the quality of the instructional experience including psychological and curricular aspects.

Four environmental factors also consistently affect learning:

the educationally-stimulating, psychological climates of
6) the home,
7) classroom social group, and 8) the peer group outside school; and
9) minimal leisure-time television viewing.

The first five aspects of student aptitude and instruction are prominent in the educational models of Benjamin S. Bloom, John B. Carroll, Robert Glaser, and others. Each appears necessary for learning in school; without at least a small amount of each, the student can learn little. Large amounts of instruction and high degrees of ability, for example, may count for little if students are unmotivated or instruction is unsuitable.

These five essential factors, however, are only partly alterable by educators since, for example, the curriculum in terms of lengths of time devoted to various subjects and activities is partly determined by diverse economic, political, and social forces. Ability and motivation, moreover, are influenced by parents, by prior learning, and the students themselves. Thus educators are unlikely to raise achievement substantially by their own efforts alone.

Of the remaining factors--the psychological climate of the classroom group; enduring affection and academic stimulation from adults at home; and an out-of-school peer group with learning interests, goals, and activities--influence learning in two ways: Students learn from them directly; and these factors indirectly benefit learning by raising student ability, motivation, and responsiveness to instruction. In addition, about ten (not the more typical 30) weekly hours of television viewing seem optimal for learning, perhaps because more television time displaces homework and other educationally-constructive activities outside school.

The major causal influences flow from aptitudes, instruction, and the psychological environment to learning. In addition, however, these factors also influence one another, and are also influenced in turn by how much students learn. since those who begin well learn faster.

Other social factors influence learning in school but are less directly linked to academic learning. For example, class size, financial expenditures per student, and private governance (independent or sectarian in contrast to public control of schools) weakly correlate with learning, especially if the initial abilities of students are considered. Thus, improvements in the more direct and more alterable factors hold the best hope for increasing educational productivity.

Thus, in my view, school and district economic, political, and sociological characteristics and conditions are less relevant to learning because their influences are less alterable, direct, and observable. They are not substitutes for the nine factors, but more distant forces that can support or interfere with them.

More and less productive classes, moreover, may be expected in the same school; and it is somewhat misleading to characterize a whole school or district as effective--just as it is less accurate to characterize an optimal condition of plant growth as the average annual rate of rainfall in a state or farm than the amount of rain and irrigation that reaches the roots of a single plant in a given time period.

The educational productivity theory itself is admittedly over-simplified because learning is clearly affected by school and district characteristics as well as many economic, sociological and political forces at the school, community, state, and national levels. Yet these characteristics and forces--such as the sex, ethnicity, and socioeconomic status of the student, the size and expenditure levels of schools and districts, and their political and sociological organization--are less alterable in a democratic. pluralistic society; are less consistently and powerfully linked to learning; and appear to operate mainly through the nine factors in the determination of achievement. Thus, I offer our theory not as a threat to those who see the efficacy of other factors but as a friendly, collegial invitation to demonstrate their effects on the nine factors or directly on learning.

Methods of Research

Since our concern was productivity, we hoped that our own research would efficiently capitalize on previous inquiry; and, under the support of the National Institute of Education and the National Science Foundation, our team of investigators started by compiling reviews of the 1970s on the productive factors in learning. Next, quantitative syntheses of studies of productive factors were conducted; syntheses of several thousand investigations were compiled (see Walberg, 1984, for a more detailed account). Case studies of Japanese and American classes were carried out to compare educational productivity in the two countries. The productive factors were further probed for their significance in promoting learning in three large sets of statistical data on elementary and high school students--the National Assessment of Educational Progress, High School and Beyond, and the surveys of the International Association for the Evaluation of Educational Achievement.

Collectively the various studies suggest that the nine factors are powerful and consistent in influencing learning. Syntheses of about 2,575 studies suggest that these generalizable factors are the chief influences on cognitive, affective, and behavioral learning. Many aspects of these factors can be altered or influenced by educators.

The first five essential factors appear to substitute, compensate, or trade-off for one another at diminishing rates of return. Immense quantities of time, for example, may be required for a moderate amount of learning if motivation, ability, or instructional quality is minimal. Thus, no single essential factor overwhelms the others; all appear important.

Although the other factors are consistent statistically- or experimentally-controlled correlates of academic learning, they may directly supplement as well as indirectly influence the essential classroom factors. In either case, the powerful influences of out-of-school factors especially the home environment must be considered.

For example, the 12 years of 180 6-hour days in elementary and secondary school add up to only about 13 percent of the waking, potentially-educative time during the first 18 years of life. If more of the 87 percent of the student's waking time nominally under the control of parents that is spent outside school were to be spent in academically-stimulating conditions in the home and peer group, then the total amount of the student's total learning time might be dramatically raised beyond the 13 percent of the time in conventional American schools.

For instance, the average of 28 hours a week spent viewing television by high school students might usefully be added to the mere 4 or 5 weekly hours of homework (Walberg and Shanahan, 1983). Europeans and Japanese believe homework helps learning; empirical results of American research support their belief.

The numerical results of syntheses of the effects in several thousand studies of academic learning conducted during the past half century. Interested readers and those who wish technical details may examine the findings and methods reported in the compilations of these syntheses (cited in the references in Walberg, 1984, which in turn, contain references to the original studies. (In several instances, separate estimates of correlations and effects are available for science and mathematics because the National Science Foundation awarded grants for special synthesis projects on these two subjects. The tables contain both effects and correlations, and the correlations assume a one- standard deviation rise in the independent variable.)

Sample survey items and descriptions of sets of items from High School and Beyond, the National Assessment of Educational Progress, and the International Association for the Evaluation of Educational Achievement provide national and international baselines in various years, and theyshould be considered as candidates on this ground alone. Better items and direct observations can also be developed and used.

Beyond Academic Achievement

If education proceeds by fads rather than cumulative research, it will fail to make the great advances in productivity that have characterized agriculture and industry in this century. It may be argued, however, that education is a complex subject and cannot be reduced to a few external benefits or measures of outcomes. This argument also applies to any enterprise: The desirability of an automobile cannot be reduced to numbers on its speed and power; bushels of corn per acre need to be considered in the light of percentage of protein per unit weight, predicted prices, fuel requirements, human labor, and the like.

Similarly, better nationally-calibrated measures of achievement including facts and "higher-order skills" in English, mathematics, science, civics, history, foreign languages, art, and music and what produces them would be a great accomplishment. But they would hardly suffice and may be misleading.

A synthesis of the relation of conventionally-measured educational outcomes and adult success shows their slight association (Samson and others, 1982). Thirty-three post-1949 studies of the college and professional-school grades of liberal arts and business graduates, nurses, physicians, engineers, civil servants, teachers, and other groups show an average correlation of .155 of these educational outcomes with life-success indicators such as income; self-rated happiness; work performance and output indexes; and self-, peer-, and supervisor-ratings of occupational effectiveness. Thus, only about 2.4 percent of the variance in these indicators of adult success was predictable from grades given by professors.

These results should challenge educators and researchers to seek a balance between continuing autonomy, motivation, responsibility, and skills to learn new tasks as an individual or group member on one hand and mastery of teacher-chosen, textbook knowledge measured on conventional tests that may soon be obsolete or forgotten on the other. Researchers need to think again about how civic virtue, perseverance, will power, cooperation, entrepreneurship and the like that are no longer in the current psychological lexicon might be measured and encouraged.

One clue comes from old studies of open education, in which teachers and students negotiated contractual terms about what students would learn. Open educators tried to encourage educational outcomes that reflect teacher, parent, student, and school board goals such as cooperation, critical thinking, self reliance, constructive attitudes, life-long learning, and other objectives seldom considered by psychometrists. Raven's (1981) summary of surveys in Western countries including England and the United States, shows that, when given a choice, educators, parents, and students rank these goals above test scores and high marks.

Hedges, Giaconia, and Gage (1981) synthesized 153 studies of open education including 90 dissertations. The average effect was near zero for achievement, locus of control, self concept, and anxiety (which suggests no difference between open and control classes on these criteria); about .2 for psychological adjustment, attitude towards schools and teachers, curiosity, and general mental ability; and about a moderate .3 for cooperativeness, creativity, and independence. Thus, students in open classes do slightly or no worse in standardized achievement and slightly to moderately better on several outcomes that educators, parents, and students hold to be of great value. Thus, this recently-synthesized old research shows the value non-standard outcomes and demonstrates that conventional measures do not necessarily enhance or trade-off against unconventional accomplishments.

Another precedent for non-conventional measurement is the current effort beginning under the sponsorship of the Swedish Ministery of Education. Sweden is fortunate in having a longitudinal sample first measured in 1961 of people born on the 5th, 15th, and 25th of all months of 1948, who are now nearing 40 years of age. Harnqvist (1984) is beginning an internationally important series of studies of this sample to discover how early school and other experiences influence adult knowledge and attitudes. Of about 120 adult characteristics, 71 percent have shown significant partial correlations with amount of education, 30 percent with measured intelligence, and 20 with social background.

Other things being equal, Swedish adults with more education, for example, more often reported that their jobs provided them with new knowledge and more influence on determining their working conditions. Amount of education was positively associated with "cultural" activities such as going to theaters and concerts, and negatively correlated with "entertainment" through weekly magazines, television, and sports events. More highly educated men reported higher skills in cooking and lower skills in repairing a car; and having better information on appealing decisions and less about seeking economic support from society. In the interviews, more educated men and women used more words, more different words, a greater percentage of words with more than 10 letters; and they required fewer interviewer interventions to complete their responses.

Harnqvist's is a pioneering longitudinal investigation. Since the effects of education may not turn up immediately, except on knowledge tests alone and less well on other measures, we need more such long-term studies that relate adult characteristics to educative experiences and activities within and outside school. We are fortunate, indeed, to have Harnqvist's contribution that shows the influence of amounts of education; but we would like to have more specific measurements of early and later accomplishments such as hobbies and prizes won, and experiences such as courses taken, homework hours, books read at leisure, family activities, trips abroad, military service, and the like. These would enable us to relate early characteristics and experience to later achievements and attitudes.

Conclusion

The present seems a time for great opportunity in educational reform and research in education. Agriculture, engineering, and medicine made great strides in improving human welfare as doubts arose about traditional, natural, and mystical practices, as the widened measurement of results intensified, as experimental findings were synthesized, and as their theoretical and practical implications were coordinated and vigorously implemented and evaluated.

Education is no less open to humanistic and scientific inquiry and no lower in priority since half the workers in modern nations are in knowledge industries, and the value of investments in people is now more apparent than ever (Walberg, 1983). Although it is possible to find fault with federal statistics on education, the last decade or two has been a period of quiet but significant accomplishments; and larger amounts of valuable data are being accumulated.

Recently the National Research Council's Committee on Indicators of Precollege Science and Mathematics Education issued a report calling for the national measurement and tracking of the many of the same productivity factors and outcomes discussed above (Raizen and Jones, 1985). The U.S. Department of Education, working with 16 education organizations, has already developed a plan for systematically collecting outcome, process, and context data and issued its first report, Indicators of Education Status and Trends (1985). This contains a series of data, presented both in tables and graphs over time, showing the course of education measures over several decades.

Both reports continue our tradition of collecting enrollment and spending data; but go beyond it in recommending (in the case of the NRC report) and displaying (in the DE report) changes in test scores, international comparisons of achievement, remedial college course enrollments, class sizes, verbal abilities of the teaching force, public-opinion ratings of schools, and state-required curriculum units. These reports give us hope that we may reach consensus on extending measures of learning and of the productive factors that bear upon it, and that a national bureau might be founded to collect, coordinate, calibrate, archive, analyze, synthesize, and make available the data that is needed to improve educational productivity.

References

Alonso, William, and Starr, Paul (1985) A Nation of Number Watchers. Wilson Quarterly, 9 (3), 93-123

Carroll, J. B. The measurement of intelligence. In R. J. Sternberg (Ed.), Handbook of human intelligence. New York: Cambridge University Press, 1982.

Cooke, Charles, Ginsburg, Allen, and Smith, Marshall (1985) The Sorry State of Education Statistics. Madison: Wisconsin Center for Educational Research.

Feistritzer, C. Emily. (1985) Cheating Our Children. Washington, DC: National Center for Educational Information.

Harnqvist, Kjell (1984) An Empirical Study of Long-Term Effects of Education. Tel Aviv, Israel: Paper presented at the First International Conference on Education in the 1990s.

Hedges, L. V., Giaconia, R. M., & Gage, N. L. Meta-analysis of the effects of open and traditional instruction. Stanford, CA.: Stanford University Program on Teaching Effectiveness, 1981.

Horn, Elizabeth, and Walberg, Herbert J. (1984) Achievement and interest as functions of quality and level of instruction. Journal of Educational Research, 77, 227-232.

Indicators of Education Status and Trends (1985) Washington, DC: U. S. Department of Education.

Leontief, Wassily, Duchin, Faye, and Szyld, Daniel B. (1985) New Approaches to Economic Analysis. Science, 112, 419-427.

Raizen, Senta A. and Jones, Lyle V. (1985) Indicators of Precollege Education in Science and Mathematics: A Preliminary Review. Washington, DC: National Academy Press.

Raven, J. The most important problem in education is to come to terms with values. Oxford Review of Education, 1981, 7, 253- 272.

Samson, G. E., Graue, M. E., Weinstein, T., & Walberg, H. J. (1984) Academic and occupational performance: A quantitative synthesis. American Educational Research Journal, 21, 311- 21.

Thurstone, L. L. The relation between learning time and length of task, Psychological Bulletin, 1930, 37, 44-53.

U. S. Commission on Excellence in Education (1983) A Nation at Risk, Washington, DC: U. S. Government Printing Office.

Walberg, Herbert J. (1984) Educational Expenditures and Student Achievement. New Orleans, Louisiana: Paper presented at the annual meetings of the American Educational Research Association.

Walberg, Herbert J. (1984) Improving the Productivity of America's Schools. Educational Leadership, 41 (8), 19-30.

Walberg, Herbert J. National Abilities and Economic Growth. New York, NY: Paper presented at the annual meetings of the American Association for the Advancement of Science.

Walberg, Herbert. J. Scientific literacy and economic productivity in international perspective, Daedalus, 1983, 112, 1-28.

Walberg, Herbert J., Harnisch, Delwyn L., and Tsai, Shiow-Ling. (in press) Mathematics Productivity in Japan and Illinois. Evaluation in Education: An International Journal, 1985.

Walberg, Herbert J., and Hess, G. Alfred (1985) Chicagoans View the Public Schools Chicago: Chicago Panel on Public School Finances.

Walberg, Herbert J., Paschal, Rosanne A., and Weinstein, Thomas (1985) Homework's Powerful Effects on Learning. Educational Leadership, 42 (7), 76-78.

Walberg, Herbert J., and Rasher, Sue Pinzur (1979) Achievement in the 50 States. In Herbert J. Walberg (Ed.), Educational Environments and Effects. Berkeley, CA.: McCutchan Publishing.

Walberg, Herbert J., and Shanahan, Timothy (1983) High school effects on individual students. Educational Researcher, 12 (7), 4-9.

Appendix

This appendix contains several tables that illustrate the magnitudes of effects of productivity factors on achievement revealed by quantitative syntheses carried out by a number of investigators in Australia, Canada, and the United States during the past decade. In addition, operational representations of the factors and sample items from re-analyses of the National Assessment of Educational Progress, High School and Beyond, and the first mathematics survey of International Association for the Evaluation of Educational Achievement are given (Horn and Walberg, 1982; Walberg and Shanahan, 1983; and Walberg, Harnisch, and Tsai, 1984).

Table 1

Influences of Aptitudes on Learning

Aptitude	Effect	Size
Ability		
IQ	.71	XXXXXXX
IQ (Science)	.48	XXXXX
Development		
Piagetian Stage	.47	XXXXX
Pia. Stage (Science)	.40	XXXX
Motivation		
Motivation	.34	XXX
Self-Concept	.18	XX

Note: The X symbols represent the sizes of the correlation coefficients in numbers of tenths.

Table 2
Instructional Quality and Time Effects on Learning

Method	Effect	Size
Reinforcement	1.17	XXXXXXXXXXXX
Acceleration	1.00	XXXXXXXXXX
Reading Training	.97	XXXXXXXXXX
Cues and Feedback	.97	XXXXXXXXXX
Science Mastery	.81	XXXXXXXX
Cooperative Programs	.76	XXXXXXXX
Reading Experiments	.60	XXXXXX
Personalized Instruc.	.57	XXXXXX
Adaptive Instruc.	.45	XXXXX
Tutoring	.40	XXXX
Individualized. Science	.35	XXXX
Higher-Order Questions	.34	XXX
Diagnostic Prescription	.33	XXX
Individualized Instruc.	.32	XXX
Individualized Math.	.32	XXX
New Science Curricula	.31	XXX
Teacher Expectation	.28	XXX
Computer-Assis. Instruc.	.24	XX
Sequenced Lessons	.24	XX
Advanced Organizers	.23	XX
New Math. Curricula	.18	XX
Inquiry Biology	.16	XX
Homogeneous Groups	.10	X
Programmed Instruc.	-.03	-.
Class Size	-.09	-X.
Mainstreaming	-.12	-X.
Instructional Time	.38	XXXX

Note: The X symbols represent the sizes of effects in tenths of standard deviations.

Table 3

Home, Peer, Class Morale and Media Effects

Method	Effect	
Graded Homework	.79	XXXXXXXX
Class Morale	.60	XXXXXX
Home Interventions	.50	XXXXX
Home Environment	.37	XXXX
Assigned Homework	.28	XXX
Socioecon. Status	.25	XXX
Peer Group	.24	XX
Television	-.05	X.

Note: The X symbols represent the sizes of effects in tenths of standard deviations or correlations.

APPENDIX NAEP

Variable Descriptions and Sample Characteristics

Operational Definition, Internal Consistency, Sample Paraphrased Items, Scoring, Percent of Sample in Each Category (when applicable).

Based on Sample of 2,294 Students.

Achievement	Fifty-five items assessing student achievement in five content categories and four cognitive-process levels. Alpha internal consistency reliability = .92.
Interests	Three self-report items probing student willingness to study mathematics not part of a classroom assignment. Alpha internal consistency reliability = .47. "How often did you work ahead in your mathematics book?" "How often did you do mathematics problems that were not assigned?" "How often did you study mathematics topics that were not in the textbook?" Coded: 3 = often, 2 = sometimes, 1 = never, blank = no response or missing.
SES	Highest amount of either parent's education. Coded: 1 = Not a high school graduate (15.6%) 2 = Graduated high school (34.8%) 3 = Post high school (44.9%) blank = Unknown or missing (4.7%)
Traditional instruction	Two items on traditional instructional methods. Alpha internal consistency reliability = .55. "How often has each of the following been used in the courses you are taking this year?" 1. Listening to the teacher's lecture 2. Studying from textbooks Coded: 3 = frequently, 2 = fairly often, 1 = seldom or never, blank = no response or missing.
Home environment	Twelve items on home characteristics. Alpha internal consistency reliability = .66. "Which of the following do you have in your home?" 1. Newspaper received regularly 2. Magazine received regularly 3. More than 25 books 4. Encyclopedia 5. Dictionary 6. Record player 7. Tape recorder or cassette player 8. Typewriter 9. Vacuum cleaner 10. Electric dishwasher 11. Two or more cars or trucks that run 12. English spoken most often at home Coded: 0 = do not have, 1 = have, blank = no response or missing.
Sex	Coded: 1 = female (51.5%), 0 = male (48.5%) Coded: 1 = white (81.8%), 0 = not white Coded: 1 = black (12.7%), 0 = not black Coded: 1 = Spanish heritage (4.8%), 0 = not Spanish heritage
Student-centered instruction	Two items on instructional methods emphasizing student participation. Alpha internal consistency reliability = .43. "How often has each of the following been used in the courses you are taking this year?" 1. Participations in student-centered discussions 2. Having individualized instruction in small groups or one-to-one with a teacher Coded: 3 = frequently, 2 = fairly often, 1 = seldom or never, blank = no response or missing.
Stimulation	Four items indicating frequency of course-related activities. Alpha internal consistency reliability = .46. "How often has each of the following been used in the courses you are taking this year?" 1. Working on a project or in a laboratory 2. Writing essays, themes, poetry, stories 3. Going on field trips 4. Library or media center assignments Coded: 3 = frequently, 2 = fairly often, 1 = seldom or never, blank = no response or missing.
Number of math courses	Seven items indicating mathematics course taken or completed; summed to indicate number of years mathematics was studied. "Which of the following mathematics courses have you studied?" 1. General, business, or consumer mathematics (38.3% studies one year) 2. Pre-algebra (36.0%) 3. First year algebra (67.3%) 4. Geometry (43.3%) 5. Second year algebra (28.7%) 6. Trigonometry (6.7%) 7. Pre-calculus/calculus (2.2%) Coded: 0 = not studied, .25 = studied less than one half of school year, .5 studied one half of school year, 1 = studied about 1 school year, blank = no response or missing.
Highest course	Indicates highest level mathematics course taken for at least one half of school year. Coded in order of course difficulty: 1 = General, business, or consumer mathematics (10.9%) 2 = Pre-algebra (7.2%) 3 = First year algebra (17.3%) 4 = Geometry (14.0%) 5 = Second year algebra (23.8%) 6 = Trigonometry (9.7%) 7 = Pre-calculus/calculus (4.0%) Blank = no response or missing (13.1%)
Homework	Amount of time spent doing homework last night. Coded: 3 = more than two hours (8%) 2 = between one and two hours (18%) 1 = less than one hour (20%) 0 = did not do homework or no homework assigned (40%) Blank = no response or missing (14%)
TV	Amount of time spent watching TV last night 8 = six or more hours (4.0%) 7 = five hours (3.6%) 6 = four hours (9.2%) 5 = three hours (11.2%) 4 = two hours (14.9%) 3 = one hour (8.0%) 2 = less than one hour (15.7%) 1 = none (18.6%) Blank = no response or missing (14.9%)

BEST COPY AVAILABLE

TABLE I

Variable Descriptions and Numbers of Items, Reliabilities, and Univariate Statistics

HSB

Variable	Descriptions or Sample Content of Items; Percentages for Categorical Items	Number of Items	Reliability	Mean	Standard Deviation
Achievement					
Vocabulary 1	Multiple-choice test	15	—	49.52	10.05
Vocabulary 2	Multiple-choice test	12	—	49.56	10.06
Reading	Multiple-choice test	20	—	49.37	10.12
Mathematics 1	Multiple-choice test	25	—	49.38	10.13
Mathematics 2	Multiple-choice test	8	—	49.65	9.98
Age	In years	1	—	17.47	6.46
Motivation					
Aspiration	Occupational goals age 30	1	—	8.47	4.39
Work orientation	(HSB scored composite)	3	—	.01	.69
	Self-satisfaction	2	.33	.90	.24
Adjustment Problems	Discipline problems, cutting, and suspension	3	.40	5.30	.80
Control locus	(HSB scored composite)	4	—	.01	.69
Motivation	Items on English and mathematics interest and usefulness, interest and hard work	11	.62	5.84	1.15
Law trouble	Serious trouble; Yes = 3.9%	1	—	1.96	.19
Quality of Instruction					
Quality	Ratings of quality of instruction such as good teaching, academic emphasis, school reputation, teacher interest in students, and instructional qualities	10	.59	5.19	.84
Quantity of Instruction					
Quantity	Academic courses completed in English, mathematics, French, German, Spanish, history and science	10	.71	4.84	.92
School Environment					
Facilities	Ratings of school building and library	2	.80	5.51	1.43
Discipline	Ratings of effectiveness and fairness of school	2	.89	4.83	1.50
Extracurricular activities	Student participation in school sports, clubs, band, and debate	15	.67	19.29	3.51
Peer Environment					
Peer	Grades of friends, their school interest in classes and college, and regular school attendance	4	.67	7.08	1.15
Home Environment					
Parent interest	Parental monitoring and interest in school, work, and career plans	9	.34	3.48	.81
Home facilities	Place to study, daily newspaper, enclyclopedia, and electric dishwasher	8	.52	14.01	1.53
Mother work	Mother working before and during elementary and high school	3	.78	6.12	2.02
Homework	Hours per week spent on homework	1	—	4.42	1.41
Age first worked	For pay	1	—	4.52	2.31
Worked last week	In hours	1	—	1.38	.48
Hours currently worked per week	In hours			3.77	1.66
Hours worked per week during previous school year	In hours			4.80	1.92
Socioeconomic status (SES)	SES composite scale (HSB scored)	5	—	−.09	.75
Media Exposure					
Television	Hours watched per day	1	—	4.43	1.70
Miscellaneous					
Handicaps	Visual, hearing, speech, learning, and health handicaps	5	.32	8.70	.60
Physically unattractive	Yes = 11.8	1	—	1.88	.31
Male	Yes = 47.9	1	—	1.48	.49
White	Yes = 75.5	1	—	.75	.43
Spanish	Yes = 11.2	1	—	.11	.32
Asian	Yes = 1.3	1	—	.01	.11
Black	Yes = 14.0	1	—	.14	.35
Alternative public schools	Yes = 3.2	1	—	.03	.18
Catholic	Yes = 9.5	1	—	.10	.29
Elite private	Yes = 1.1	1	—	.01	.11
Other private	Yes = 2.0	1	—	.02	.14

BEST COPY AVAILABLE

Table 1

Variable Description for IEA

Variable	Operational Definition, Scoring, and Sample Items
Mathematics Achievement	70 items assessing student mathematics in both content (mathematics subject matter) and cognitive-process levels (e.g., particular skills, abilities, and knowledge, etc.) Range of internal-consistency reliability (Kuder-Richardson 20) among 12 countries is.73 to .93, where the United States is .84. "The result of an operation on the numbers 9 and 18 is 27. In this operation, the number 27 is (a) product; (b) sum; (c) quotient; (d) difference, (e) average." "Four persons whose names begin with different letters are placed in a row, side by side. What is the probability that they will be placed in alphabetical order from left to right? (a) 1/120; (b) 1/24 (correct); (c) 1/12; (d) 1/6; (e) 1/4."
Male	Coded: 1-male, 0-female
Socioeconomic Status	Highest amount of either parent's education; code is indicated by the number of completed years.
Scientific Background	To indicate father's occupation: 1-scientific, 0-nonscientific
Father's Occupation	Coded: 1-higher professional and technical occupations 2-farm proprietors and farm laborers 3-subprofessional technical, small worker-proprietor (non-farm), clerical, and sales 4-manual workers (non-farm)
Mother's Employment	"Is your mother presently working?" Coded: 0-not working, 1-part time, 2-full time
Highest Mathematics Course Taken	"Indicate the highest level of mathematics courses that you have taken recently: (coded:) 1-arithmetic or general mathematics 2-algebra 3-geometry 4-trigonometry 5-advanced mathetics (calculus, etc.)"

Table 1 (page 2 of 2)

Variable	Operational Definition, Scoring, and Sample Items
Number of Students in Math Class	"Indicate the approximate number of students in your present or most recent mathematics course: (coded:) 1-under 10 5-25 through 29 2-10 through 14 6-30 through 34 3-15 through 19 7-35 through 39 4-20 through 24 8-40 or more"
Periods of Math per Week	"In your mathematics class, how many periods do you have each week? (coded): 1-1 or 2 5-9 or 10 2-3 or 4 6-11 or 12 3-5 or 6 7-13 or more" 4-7 or 8
Hours of Homework	"Indicate the amount of hours that students usually devote to homework each week."
Extra Mathematics Activities	"Have you been a member of any mathematics club, or attended special lectures or courses on mathematics? (coded:) 1-yes, 0-no"
Interest in Mathematics	Ten items are included to measure the level of interest in mathematics; e.g., "Wishes to take additional math couses (coded:) 1-yes, 0-no" "Which two school subjects have you liked most? (coded:) 1-mathematics, 0-others
Attitude Toward	Eleven items form a scale to ascertain the student's disposition toward school life, e.g., "I find school interesting and challenging. (coded:) 2-agree, 0-disagree" "I am bored most of the time in school. (coded:) 2-disagree, 0-agree"

Toward Achieving Educational Excellence for All Students

Margaret C. Wang
Learning Research and Development Center
University of Pittsburgh

Working to provide school learning environments that lead to educational success for all students has been a staple of reform efforts throughout the history of public education in the United States. In recent years, this goal has been conjoined with that of improving schools' capabilities to effectively serve increasingly diverse student populations, particularly students from economically, culturally, and language disadvantaged backgrounds and other academically-at-risk students. Public, judicial, and legislative movements calling attention to both equal access and equal chances to achieve schooling success have been a key factor in the intersecting of these two goals. The enactment of the Education for All Handicapped Children Act of 1975 (PL 94-142) is a prime example of such efforts.

Significant progress has been made, especially during the past decade, in providing equal access to free and appropriate schooling for all students. Now that virtually all school-aged children and young adults attend school (Plisko, 1984), ensuring equal opportunities for schooling success through quality education has become a national priority (e.g., National Coalition of Advocates for Students, 1985; National Commission on Excellence in Education, 1983; Twentieth Century Fund, 1983; U. S. House of Representatives, 1983). The current push for educational excellence recognizes that the critical goal of merging human resources and talents with technological and economic growth cannot be achieved without a concerted effort to curtail the rising number of students with poor prognoses for academic success while also increasing educational efficacy and productivity for all students.

Thus, a number of convergent realities challenge us to bring about improved education for all students. These include public sentiment, legislative and regulatory developments, the need to invest in human capital, technological innovations, and demographic changes. Greater-than-usual educational support is likely to be required for accommodating the expanding literacy requirements that accompany rapid technological growth in our advancing society as well as the anticipated continuing trend of increasing proportions of the school-aged population from culturally, economically, and language disadvantaged backgrounds. In responding to this challenge, school improvement efforts can draw from parallel developments in research on learning and effective schooling, advances in educational technology, and the development and implementation of innovative programs.

In this context, this paper has two main purposes. The first is to discuss the implications of findings from the past decade of educational reform aimed at improving the quality of schooling — particularly for special needs students. (Special needs students are defined here as students who are academically-at-risk and/or who receive special education or compensatory education services such as Chapter I, bilingual education, or migrant education.) The second purpose of the paper is to identify information needs and formulate recommendations to the National Center for Education Statistics (NCES) regarding specific ways of broadening the current data base to enhance its relevance for description, analysis, and improvement of the conditions and quality of schooling for all students.

RECENT DEVELOPMENTS AND CURRENT SCHOOL IMPROVEMENT EFFORTS

The State of Practice

In response to the call for both equal access to schooling and educational success, the past 20 years have witnessed a proliferation of legislation and federally-supported "special" or "compensatory" remedial education programs. Actual steps to improve the educational attainment of special needs students generally have focused on creating separate programs that have been implemented, for the most part, in settings segregated from regular education students and classes. The growth of this segregated approach was especially rapid in the 1970's when, for example, students with learning problems (e.g., low achievement, disciplinary problems) were treated as "special" and relegated to separate programs implemented primarily by specialists. This approach, albeit well-intentioned, neglected to recognize the larger problem – that regular school environments had failed to accommodate the educational needs of a large number of students. On the one hand, it can be argued that special programs are positive steps reflecting a commitment to provide effective instruction for all children. However, the "set-aside" strategy on which the programs are based is driven by the fallacy that poor school adjustment and performance are attributable solely to characteristics of the child rather than to the quality of the school learning environment.

The problems faced by schools in their efforts to adequately provide for special needs students are derived from a variety of sources. These range from changes in national educational priorities to the increased focus on procedural rather than programmatic issues. Included among the specific road blocks to equal access and educational excellence for special needs students are the redeployment of critical resources (human and fiscal) from the provision of education to the administration of an inordinate number of separately funded and delivered programs; the proliferation of classifications for students with special learning needs, combined with the growing use of noninstruction-related criteria to label and classify greater and greater numbers of students in mildly handicapped categories such as learning disabled (LD); a downshift in public and financial support for school programs in both regular and special education; inadequate personnel preparation programs for instructional staff – including regular and special educators – who are expected to develop and implement effective instruction for students in the many special and compensatory education categories; and an overall lack of coordination between educational experiences in special and regular education settings.

Observations in a variety of school settings have suggested that reliance upon separate school improvement programs often leads to piecemeal remediation in segregated environments. While "pull-out" programs may be helpful for certain students (e.g., severely disabled students), such programs are more likely to have negative results, including discontinuities and interruption in the instructional-learning process for teachers and students, loss of control by school district leadership over specialized programs, and the fostering of narrow categorical attitudes and instructional programming (cf. Heller, Holtzman, & Messick, 1982; Reynolds & Wang, 1983). As further evidence of the inadequacy and arbitrariness of the segregated approach to providing instruction for special needs students, enrollments in the least well-defined categories such as LD are climbing beyond the tolerance of budget makers; the courts are taking away many of the "special" classification procedures (as in *Larry P. v. Riles*, 1972); scientists tell us that most of the diagnoses performed in special education essentially are unrelated to treatments (Reschly, in press; Ysseldyke, Algozzine, & Epps, 1983); and the monitoring of compliance with special education legal requirements tends to be more procedural than substantive. In recent years, the problems associated with the many separate programs have resulted in the emergence of new approaches to more effectively meeting the diverse needs of most, if not all, students in regular school environments.

The Setting for Change

Finding feasible and effective ways to restructure special and compensatory education programs and delivery systems clearly is a pressing current need. It is important to keep in mind, however, that this restructuring process must take place in the context of the entire educational enterprise. Special and compensatory education programs often are expected to accomplish what otherwise would be left undone, or done poorly, by regular education. If headway is to be made in the effective academic and social integration of special needs students in regular classes, regular education staff and specialized professional personnel at the federal, state, and local levels must work together to negotiate the removal of many of the present barriers. We need to achieve a healthy balance between the current preoccupation with classification of students for educational placement, efforts to identify specific instructional needs, and the implementation and refinement of available instructional solutions.

Advances in Research and Innovative Program Development

The history of educational reform generally has been characterized by patterns of parallel developments in psychological theory of learning, technical advances in instructional practices, and socioeconomic and political mandates of the time. Likewise, the beginning phase of the current new wave of developments is marked by significant progress in research on learning and effective teaching, an intense motivation to improve schools' capabilities for effectively responding to student diversity, innovative program development and the implementation of school improvement efforts, and the sociopolitical mandate to maximize the chances of schooling success for all students.

A number of alternative interventions have been developed and tested for integrating special needs students in regular classes. Recent research literature as well as the reports of several prominent groups (e.g., Cantalician Foundation, 1983; Heller et al., 1982; Mayor's Commission on Special Education, 1985) point to the potential of these interventions. Some of the specific recommendations that have emanated from recent reports are the full academic and social integration of special needs students (handicapped students and other children at the margin) in regular classes and schools; the provision of appropriate educational experiences based on learning needs rather than on rigid classifications, labels, and placements; and the restructuring of regular education to more effectively accommodate all individual students, regular and special needs students alike. Researchers, policymakers, and practitioners all have become lobbyists for the delivery of compensatory and special education services in regular classes; in doing so, they have espoused the educational vision of ensuring quality schooling services for the increasingly diverse student populations our nation's schools are challenged to serve (Heller et al., 1982; National Coalition of Advocates for Students, 1985; U.S. House of Representatives, 1983).

Implementation and Fiscal Barriers

In spite of recent research findings and experience with innovative program development and implementation that point to the feasibility and efficacy of integrating special needs students in regular classes, special education programs and a wide variety of compensatory education programs continue to operate as separate systems (often more accurately characterized as "nonsystems"). In many cases, overlapping separate services are provided for the same students. Implementation of an integrated approach to improving educational conditions in this area will require fundamental programmatic, organizational, and funding policy reforms. A first step in this direction would be the establishment of an open, experimental period, during which regular, special, and compensatory education could be combined to encourage innovative development aimed at providing improved and integrated educational services along a full continuum -- including supplementary aids and pre-referral services in regular classes. In local schools, leadership should be encouraged for experimentation and for evaluating the effectiveness of a variety of educational approaches in solving the widespread persistent problem of how to achieve more productive learning for all students. Attention must be directed to putting into operation the most promising ideas and practices and, at the same time, making the necessary policy changes.

Information Needs

If the implementation of an open system for educational restructuring is to occur with a high level of precision and credibility, efforts during the next decade must include the development of a data base on a variety of alternative programs of educational excellence for all students. This improvement orientation dictates gathering the kinds of information that further understanding and specification of what constitutes effectiveness (indicators of efficacy); the conditions that influence effectiveness (e.g., program features and classroom environments); and the features of cost-effective, alternative programs and practices, particularly programs and practices directed at students with poor prognoses for educational success.

An overriding design concern in the task of gathering information on the conditions and impact of educational programs is the extent to which the resulting data base will be useful to researchers, educators, policymakers, and parents in their choice of avenues for improving schools' capabilities to become increasingly more effective in maximizing the chances of schooling success for all students. In this context, information-gathering agencies like NCES can play a critical leadership role in turning around the current preoccupation with collecting data for trend analysis and forecasting purposes only. The focus more appropriately should be shifted to a data-based approach whereby information is used to guide the formulation of visionary educational improvement goals and agendas for supporting futures research.

A major limitation of the current NCES data base is its lack of utility for meeting the information needs of school improvement efforts. As noted by Ravitch (1983), it is designed almost exclusively to gather data on the socioeconomics, rather than the quality, of education. Like the extant data sets being compiled by other federal, state, and local educational agencies, the NCES data base can be characterized as predominantly "status" in nature. It consists mainly of information such as enrollment trends, cost per pupil, student achievement as measured by standardized tests, and teacher-student ratios and other status information derived primarily from easily accessible quantifiable data. Status-oriented information is admittedly useful for conducting trend analysis aimed at describing the nation's educational enterprise from the socioeconomic and/or political perspectives. However, these data provide little information for informed decision making on the quality of education — that is, the crucial conditions and instructional practices for creating school learning environments that facilitate educational effectiveness.

RECOMMENDATIONS FOR BROADENING THE NCES DATA BASE

The discussion in this section centers on three topics. They are (a) the rationale and research bases for broadening the NCES data base to include information on program features, implementation conditions, and a wide array of program efficacy indicators; (b) the specific types of data that should be included; and (c) the implications for using the NCES data base to formulate, monitor, and evaluate school improvement efforts.

Rationale and Research Bases

In making recommendations to NCES, two major areas of concern are addressed. The first is the need for information on the learning environment (where, how, and the conditions under which instruction and learning take place). The second area of concern is the need for information on a variety of outcomes of effective schooling, particularly what students learn beyond the basic skills as measured by achievement tests (e.g., the quality of students' functioning in and outside of the school learning environment, students' ability to learn on their own and from others, students' perceptions of self-competence). The rationale and research bases for addressing these concerns are discussed below.

Features and the Efficacy of School Learning Environments

The design and implementation of school learning environments that enable each student to achieve desired educational outcomes are at the core of effective schooling. The basic premise here is that, insofar as learning is a function of a student's response to the school learning environment, instruction is the intentional manipulation of the learning environment to facilitate appropriate student responses. A major complicating factor in this purposeful design and use of learning environments is the diverse requirements of individual students for achieving given outcomes. Thus, the task of improving the quality of schooling is twofold. It involves increasing schools' capabilities to effectively accommodate the unique learning needs of individual students, while also providing instructional interventions that enhance each student's ability to respond effectively to schooling and thereby to eventually attain intended outcomes. Accomplishing this dual-natured task is a continuing challenge for educators.

Responses to this challenge have included a variety of research and development efforts with significant implications for the design and implementation of educational practices that enable students, including students with special learning needs, to maximize their chances for learning success. In fact, the development of practicable educational interventions that provide greater-than-usual educational support to accommodate the learning needs of individual students has been the hallmark of effective schooling (cf. Brandt, 1985). Four recently completed research integration studies are discussed below to provide an overview of these developments. All four studies were designed to identify critical features of widely-implemented educational interventions or approaches, as well as investigate the relationship of the program features to a variety of desired student learning processes and outcomes. When considered collectively, findings from the four studies represent a comprehensive analysis of the state of the art and the state of practice in topics related to schooling and student diversity. For the specific purposes of this paper, they serve to illustrate in particular the research base for the kinds of information on the features and efficacy of school learning environments that should be included in the NCES data base.

The first study involved the compilation and summary of findings from over 2,500 studies of educational effects on learning (Walberg, 1984). It was conducted to identify major causal influences on educational productivity. The second and third studies were designed to identify characteristic features of programs that provide for student differences. One of these studies was a quantitative synthesis of studies of features and outcomes of instructional programs aimed at adapting to student differences (Waxman, Wang, Anderson, & Walberg, 1985). The other was a quantitative synthesis that focused on the features and efficacy of mainstreaming, or the integration of handicapped students in regular classes (Wang, Birch, Anderson, & Reynolds, 1985). The final study was a large-scale, classroom observation study of program features, classroom processes, and outcomes in exemplary classes of eight instructional models (Wang & Walberg, in press).

Findings from the four studies are summarized under two headings: productivity factors in learning, and characteristic features of learning environments that provide for individual differences.

Productivity Factors in Learning

Nine interrelated factors have been found to be consistently associated with student learning outcomes (Walberg, 1984). These factors fall into three categories: student characteristics (aptitude); instruction; and environment. The three factors in the student characteristics category are ability or prior achievement, as measured by standardized tests; development, as indexed by chronological age or state of maturation; and motivation or self-concept, as indicated by personality tests or students' willingness to persevere intensively in learning tasks. The instruction category consists of two factors: the amount of time that students are engaged in learning; and the quality of instructional experiences, including psychological and curriculum aspects. The four factors in the environment category are the educational and psychological climates of the home, the classroom social group, the peer group outside of school, and the use of out-of-school time.

Overall, the major causal influences on student learning flow from student characteristics, instruction, and the environment to learning. Furthermore, the three categories of factors also influence each other and, in turn, they influence how much students learn. For example, each of the five factors in the student characteristics and instruction categories appears necessary for learning in schools; without at least a small amount of positive influence of each factor, a student learns little. Large amounts of instruction and high degrees of ability may not count for much if a student is not motivated or if instruction is unsuitable. Thus, findings from Walberg's research synthesis not only provide empirical support for examining the conditions of schooling and their impact from multiple perspectives, but the findings also suggest that such examination requires a host of information on both the quality of schooling and a wide range of outcomes.

Characteristic Features of Learning Environments That Provide for Student Differences

The past decade of instructional experimentation and innovative program development and implementation aimed at improving schools' capabilities to effectively accommodate students with diverse characteristics and learning needs has resulted in a substantial research base. A rather consistent list of salient features of programs aimed at accommodating individual differences can be derived from the extant research base. Findings from two quantitative syntheses of empirical studies reported in the literature during the past decade provide a summary analysis of this research. The first synthesis included 38 empirical studies of adaptive instruction that were published in the period from 1972 through 1982 (Waxman et al., 1985); the data base consisted of a combined sample of approximately 7,200 students. The second synthesis was designed to characterize the program design features and effects of instructional interventions for mainstreaming handicapped students in regular classes (Wang, Birch, Anderson, & Reynolds, 1985). This study was based on statistical data from 29 empirical studies of mainstreaming effects.

A common core of program features is distinguishable across the studies of adaptive instruction reviewed by Waxman et al. (1985). These features include (a) instruction based on the assessed capabilities of each student; (b) materials and procedures that permit each student to make progress in mastering instructional content at a pace suited to his or her abilities and interest; (c) periodic evaluations to inform each student of his or her progress toward skills mastery; (d) students' assumption of responsibility for diagnosing their present needs and abilities, planning individual learning activities, pursuing the planned activities, and evaluating mastery; (e) alternative activities and materials for aiding each student's acquisition of essential academic skills; (f) student choice in selecting educational goals, outcomes, and activities; and (g) students assisting one another in pursuing individual goals and cooperating to achieve group goals. Similar features also were found to be prominent in the data from studies included in the quantitative synthesis of mainstreaming program features and effects (Wang, Birch, Anderson, & Reynolds, 1985). The design features cited most frequently among the mainstreaming programs are continuous assessment, use of alternative routes and a variety of materials, individualized progress plans, student self-management, peer assistance, and instructional teaming.

The features of adaptive instruction found in the two quantitative syntheses of extant empirical studies discussed above also were noted in the results from the observation study of design features, processes, and outcomes of eight widely-used contemporary educational programs (Wang & Walberg, in press). Many of the programs included in the Wang and Walberg study are identified by descriptors such as individualized instruction, mastery learning, and adaptive education, and they incorporate program features such as cooperative learning, differentiated staffing, and computer-assisted management and instruction. A number of the programs are considered to be prototypes, and several have been widely adopted by schools (Jeter, 1980; Rhine, 1981; Talmage, 1975). The eight programs included in the study are the Adaptive Learning Environments Model (Wang, Gennari, & Waxman, 1985); the Bank Street Model (Gilkeson, Smithberg, Bowman, & Rhine, 1981); the Behavior Analysis Model (Ramp & Rhine, 1981); the Direct Instruction Model (Becker, Engelmann, Carnine, & Rhine, 1981); Individually Guided Education (Klausmeier, 1972); the Mastery Learning approach (Bloom, 1968); Team-Assisted Individualization (Slavin, 1983); and the Utah System Approach to Individualized Learning (U-SAIL) (Jeter, 1980). Classrooms that represent exemplary implementations of these programs were identified by the program developers and served as the sample pool for the study.

A major objective of this observation study was to identify (a) the specific features of adaptive instruction that are integrated into working programs with the design objective of making instructional provisions for individual differences, and (b) the kinds of classroom processes and outcomes typically associated with these programs. Data from 65 second-, third-, and fourth-grade classrooms provided information on contextual characteristics of the programs and the implementation sites, critical features of adaptive instruction as they were implemented, and the nature and patterns of classroom processes.

Analysis of the design features of the eight programs was based on two data sources: program design documents on each program, and data from classroom observations of the sample program classrooms. Results from analysis of the program design documents suggest that, although the eight programs are derived from different theoretical bases and use different instructional strategies, there are striking similarities in program features across their designs. For example, a diagnostic-prescriptive component is a core feature of all the programs included in the study. Each program emphasizes the importance of prescribing tasks that are appropriate for the learning needs of individual students. Similarly, each of the programs uses assessment procedures to determine whether students have achieved objectives and are ready to move on, or whether they need further instruction or practice. All the programs stress the need to maintain current and accurate records of each student's placement and progress – information which also is used in instructional planning.

One of the most salient findings of the analysis of program features is that all of the programs emphasize the importance of incorporating a broad range of demonstrably effective strategies and practices. No one specific set of strategies is claimed by any of the programs to be a panacea for solving all educational problems. The wide variety of strategies and practices adopted by the eight programs for accommodating student differences range from teacher-led, group instruction to student-initiated, individualized activities; from peer tutoring to student-cooperative work; and from the use of contingency contracts to student choice and scheduling of activities. Thus, each of the programs includes in its design a core of instructional practices (not unlike those cited most frequently in the effective-teaching literature) that are implemented in various ways to meet school improvement needs and goals.

Learning Processes as Outcomes of Effective Schooling

Throughout the history of formal education, improving students' ability to function as active learners who assume responsibility for the acquisition of knowledge and skills and are motivated to sustain patterns of self-directed, life-long learning has been prominently and consistently identified as a major objective of schooling. It is generally agreed that a basic goal of education is to enable individual students with diverse learning characteristics and needs to acquire those fundamental skills that facilitate continuing learning as well as positive feelings about, and confidence in, their personal capabilities for achieving schooling success. The crucial task in achieving this educational goal is to find ways of helping schools to become increasingly effective in creating learning environments that not only foster basic skills development but also prepare students to make the educational, occupational, and professional choices that each person deserves the chance to make. Each student should be empowered with the knowledge and skills required to think and to participate in and shape the socioeconomic and political worlds in which he or she lives.

Recent theoretical advances and expanded empirical bases regarding the nature of learning and instruction and the effects of innovative educational programs provide a rich data base for furthering our understanding and characterization of students' knowledge and skills acquisition and their effective functioning in the schooling process. This data base has broadened our conceptualization of the learning process and learner outcomes.

There has been a major shift in the kinds of information on student learning that are being gathered by researchers and practitioners. A preoccupation with information on achievement in basic skills acquisition gradually has given way to an emphasis on the different cognitive processes that are used by individual students to mediate the acquisition and retention of knowledge and skills. Instead of characterizing student learning solely by outcome measures, there has come to be increased recognition of the importance of analyzing the processes by which individuals learn as well as the specific ways in which variations in learning performance are related to the adoption of particular learning processes for specific tasks by individual students. More and more, learner differences are characterized in terms of the manner in which information is processed, the mental mechanics and rules that students bring to the instructional environment, the motivation and affective response tendencies involved in the acquisition and retention of knowledge, and the knowledge and competence of individual students (cf. Wang & Lindvall, 1984).

Growing research evidence suggests a wide range of variability in the ways that students acquire, organize, retain, and generate knowledge and skills. As a result, researchers and practitioners are giving increased attention to instruction that is based on the specific learning needs of individual students. These needs are identified through analysis of the processes by which students acquire and retain knowledge and skills. Rather than being viewed as static, such learner characteristics have come to be considered alterable. Concomitantly, learner characteristics are less likely to be identified through traditional tests. Instead, they are identified and described according to the manner in which students process information and the knowledge and competence they possess for specific learning tasks (cf. Glaser & Bond, 1981).

This changing conception of the individual's learning process, combined with recent developments in research on classroom processes, has many implications for the ways in which learning and instruction are examined and described. Of particular interest is the reexamination of students' role in the learning process and the relationship between students' functioning in classroom learning and learner outcomes. In this context, students are conceptualized as active information processors, interpreters, and synthesizers (e.g., Brown, 1978; Doyle, 1979; Segal, Chipman, & Glaser, 1985; Wang & Peverly, in press). Individual learners are expected not only to take greater responsibility for managing, monitoring, and evaluating their learning, but also to be instrumental in adapting the learning environment to their needs and goals (e.g., identifying and obtaining learning resources) and adjusting themselves to the demands of the learning process.

Underlying this view of the active learner role is the assumption that essentially all learning involves both external and internal adaptation. External adaptation occurs in the ideas and content that are to be learned and in the modes and forms in which content is presented to the learner. Internal adaptation takes place in the learner's mind as new content is assimilated and internal mental structures are modified to accommodate the new content. Operationalizing the view that students' ability to make adaptations in their learning process is an individual difference variable with significant relevance for schooling success requires descriptions of greater varieties of learner competencies.

Three examples of information on learning-as-adaptation are discussed here as illustrations of candidate items for broadening the NCES data base to include indicators of learning outcomes beyond standardized achievement scores.

Use of Resources

In most learning environments, students are encouraged to use a variety of resources (e.g., time, curriculum materials, instructional and management help from teachers and peers) to facilitate completion of their schooling tasks. Even in classrooms where this is not the case, successful learners have been found to seek out and use supplementary resources. For example, in situations where the emphasis is on large-group instruction and where the only form of in-class presentation is the teacher's lecture, successful students make adaptations such as seeking supplementary reading sources, discussing lesson content with fellow students, spending greater-than-usual amounts of time on particular tasks, and arranging for personal conferences with the teacher. One important difference between programs that provide for individual differences and more conventional, group-paced programs is the former's built-in provision for assisting students in making these types of adaptations (e.g., by making alternative materials available, using a variety of instructional-learning procedures, allowing varying amounts of time for individual students to learn and to receive additional information). Descriptions detailing the nature and patterns of students' use of resources to facilitate their learning are likely to provide an important data base for characterizing an aspect of student competence that is integral to effective learning.

Study Strategies and Use of Specific Lesson Materials

Different learners who study the same chapter in a textbook, listen to the same lecture, or have any other common exposure to presentation of a lesson probably use different techniques in adapting lesson materials to their individual methods of learning. One may attempt to outline the chapter while reading it. Another may first scan the chapter and formulate questions that he or she will seek to answer while reading. Still another learner may underline key sentences during the first reading and then reread the underlined sentences in a review of the chapter. Individuals develop such techniques to adapt lesson materials to their learning needs.

Many teachers typically assist students in this type of adaptation by providing instruction on study strategies, but it is likely that most students learn "how to learn" very much on their own. Data on students' use of study strategies or skills, such as various aids to the comprehension of text, specific techniques for facilitating memorization of essential content, effective problem-solving skills, and related procedures, may constitute important information for curriculum developers and teachers. This kind of data base is likely to increase their understanding and their capabilities to help students become increasingly more competent in adapting lesson materials and in making the kinds of internal adaptations that facilitate learning.

Learning Behaviors and Motivation

Much research on adjustments in classroom learning has focused on cognitive process requirements in subject-matter learning. However, the goals of effective schooling obviously include learning in areas other than subject-matter achievement. Findings from research suggest that certain classroom environments may be more or less conducive to the development of schooling outcomes that are equally as valuable as subject-matter learning. These outcomes include positive attitudes toward learning, motivation for life-long learning, independence and self-responsibility, and social and personal skills. Moreover, some student characteristics (e.g., achievement level, affective response tendencies, temperament, perceptions of task/environment demands and affects, self-concept, work habits) are more or less effective than others in response to certain task/environment demands.

A data base on such behavioral patterns would greatly increase teachers' understanding of students' functioning. The individual learner possesses a unique profile of instructionally relevant characteristics which, in their interaction with particular elements in the classroom learning environment, elicit particular learner behaviors that may or may not facilitate certain learning conditions. The roles students are expected to play or, in other words, the adjustments students are required to make in their behavior and their motivation for effective functioning in different classrooms, are likely to vary greatly. Students' functioning in classrooms that are characterized by a preponderance of teacher-directed activities involving students working as a whole class at the same pace is likely to be very different from students' functioning in classroom environments that are characterized by the predominance of flexible instructional-grouping patterns, student responsibility and initiative, the availability of a variety of materials, peer assistance, and adequate time for teachers to respond to students' requests for assistance.

Recommendations of Specific Types of Data

Two lines of specific information are suggested for inclusion in the NCES data base. They are (a) information on the school and classroom learning environments, and (b) information on student outcomes. It is important to note that the recommended foci are meant to broaden the current NCES data base, not to replace it. Status information such as achievement test scores, school demographics, and information on instructional and related service staff is viewed as being important for future planning – but at a different level. The kinds of data on the learning environment and student outcomes discussed in this paper would constitute additional components of the NCES data base that would address more directly many of the current "quality of education" concerns. These data are summarized in Tables 1 and 2.

Table 1 is a sample list of educational goals and corresponding expected student outcomes that are included in the recommended information foci – the classroom learning environment and a broad array of expected outcomes. Based on analysis of the research bases on schooling requirements and the outcomes of effective schooling that are discussed in this paper and elsewhere (cf. Good, 1985; National School Public Relations Association, 1981; Wang & Walberg, 1985; Wittrock, in press), four major

Table 1

Educational Goals and Examples of Expected Student Outcomes of Effective Schooling

Educational Goals	Examples of Expected Student Outcomes
Mastery of subject-matter content	• Mastery of the curriculum content and skills necessary for effective functioning and further learning (e.g., the conventional basic-skills subjects such as reading, math, social studies, and science; as well as learning skills such as reasoning, remembering, comprehension, problem solving, oral communication, and writing)
Acquisition of a variety of learning skills	• Ability to study and learn independently • Ability to plan and monitor one's own learning activities • Ability to obtain needed assistance from others and provide assistance to others in learning situations
Development of positive attitudes toward learning	• Enjoyment in taking part in learning activities • Viewing the receiving of help from peers and the assisting of others in their learning as positive learning experiences that are integral to the classroom learning process • Special interest in certain learning areas • Motivation to continue learning and to persist in overcoming learning difficulties
Development of positive self-perceptions	• Confidence in one's ability as a learner • Confidence in oneself as a contributing member of the school/community • Confidence in one's ability to take self-responsibility for learning and behavior • Perceptions of internal locus of control

Table 2

Examples of Features of Effective Classroom Learning Environments and Expected Student Outcomes

Features of Classroom Learning Environments	Examples of Expected Student Outcomes					
	Mastery of content and skills for effective functioning	Mastery of content and skills for further learning	Ability to study and learn independently	Ability to plan and monitor learning activities	Ability to obtain assistance from others	Enjoyment in taking part in learning activities
Instructional content that is:						
• essential to further learning	X	X				
• useful for effective functioning in school and in society at large	X	X				
• clearly specified	X	X	X	X		
• organized to facilitate learning	X	X	X			
Assessment and diagnosis that:						
• provide appropriate placement in the curricula		X	X			
• provide regular assessment of progress and feedback		X	X	X		
Learning experiences in which:						
• ample time and instructional support are provided for each student to acquire essential content	X	X				X
• disruptiveness is minimized	X	X				X
• students use effective learning strategies/study skills	X	X	X		X	
• each student is expected to and actually experiences success in achieving mastery of curriculum content, and accomplishments are reinforced	X					X
• alternative instructional strategies, student assignments, and activities are used	X		X	X		X
Management of instruction that:						
• permits each student to master many lessons through independent study			X			
• permits each student to plan his or her own learning activities			X	X		
• provides for students' self-monitoring of their progress with most lessons				X		
• permits students to play a part in selecting some learning goals and activities			X	X		X
Collaboration among students that:						
• enables students to obtain necessary help from peers	X	X		X	X	
• encourages students to provide help	X	X			X	X
• provides for collaboration in group activities	X	X	X	X	X	X

BEST COPY A

Table 2 (cont.)

Examples of Features of Effective Classroom Learning Environments and Expected Student Outcomes

Features of Classroom Learning Environments	Examples of Expected Student Outcomes						
	Viewing help-giving and help-receiving as positive experiences	Special interest in certain learning areas	Motivation for continuing learning	Confidence in one's ability as a learner	Confidence in oneself as a contributing member of the school/community	Confidence in one's ability to take self-responsibility for learning and behavior	Perceptions of internal locus of control
Instructional content that is:							
• essential to further learning		X					
• useful for effective functioning in school and in society at large		X	X				
• clearly specified				X		X	
• organized to facilitate learning				X		X	
Assessment and diagnosis that:							
• provide appropriate placement in the curricula			X	X			
• provide regular assessment of progress and feedback			X	X			
Learning experiences in which:							
• ample time and instructional support are provided for each student to acquire essential content		X	X	X		X	X
• disruptiveness is minimized			X				
• students use effective learning strategies/study skills			X	X		X	
• each student is expected to and actually experiences success in achieving mastery of curriculum content, and accomplishments are reinforced			X	X	X	X	
• alternative instructional strategies, student assignments, and activities are used		X	X	X			
Management of instruction that:							
• permits each student to master many lessons through independent study			X	X		X	X
• permits each student to plan his or her own learning activities			X	X		X	X
• provides for students' self-monitoring of their progress with most lessons		X	X	X		X	X
• permits students to play a part in selecting some learning goals and activities		X	X	X		X	X
Collaboration among students that:							
• enables students to obtain necessary help from peers	X			X		X	X
• encourages students to provide help	X	X			X		X
• provides for collaboration in group activities	X	X		X	X	X	X

BEST COPY AVAILABLE

educational goals of effective schooling have been delineated. As shown in Table 1, these goals are (a) mastery of subject-matter content, (b) acquisition of a variety of learning skills, (c) development of positive attitudes toward learning, and (d) development of positive self-perceptions.

Table 2 provides a summary list of program features that are suggested in the literature to be facilitative in fostering the variety of expected student outcomes of effective schooling. Referring to the table, for example, the feature "assessment and diagnosis that provide appropriate placement in the curricula" is associated with at least four expected student outcomes – mastery of content and skills for further learning, ability to study and learn independently, motivation for continuing learning, and confidence in one's ability as a learner.

Implications for Use of the NCES Data Base in School Improvement

As previously mentioned, the recommendation to broaden the NCES data base to include information on conditions of learning (the context of ongoing innovative school improvement efforts) as well as student outcomes data beyond standardized achievement test scores is derived from an "improvement" orientation rather than a "forecasting" orientation. The latter orientation is predominant in the design of most large-scale data bases, including those developed and maintained by NCES. The underlying assumption of this paper is that trend projections are simply extensions of the past. By contrast, the proposed improvement orientation recognizes and anticipates future changes; it integrates forecasting with strategic planning. In this context, effective data bases are those that provide the foundation for developing alternative futures scenarios and for making informed choices that strategically "create" futures.

Selected findings from the previously-cited observation study of eight contemporary educational programs (Wang & Walberg, in press) are discussed here to illustrate the potential use of data on the quality of education (e.g., data on program features and classroom processes) for planning and informed decision making. The results from two types of analyses are discussed as examples of potential data utilization – findings from a series of analyses that focused on delineating differences among the eight programs, and findings from a series of analyses of relational patterns between features and classroom processes.

Differences Among Programs

Findings based on the classroom observation data suggest some significant differences among the eight programs. For example, the Bank Street Model classes were observed to have the highest number of indicators of personal interactions with teachers. Programs with classrooms that were observed to have the highest numbers of indicators in other areas included, Mastery Learning – teachers' use of explaining and demonstrating/modeling; the Adaptive Learning Environments Model – students working on independent tasks in group settings, use of exploratory learning materials, one-to-one tutoring, and teachers encouraging student self-responsibility; the Behavior Analysis Model – responding, praising behavior, and cueing or prompting; Team-

Assisted Individualization — constructive student interactions, students working alone on independent tasks, students assisting in classroom management, and student assessment of task difficulty; and the Direct Instruction Model — small-group instruction, reading, and communicating criteria.

With respect to classroom processes among the eight programs, the Adaptive Learning Environments Model and Team-Assisted Individualization were observed to be most prominent and distinctive. Classes using these two programs had the most indicators of adaptive instruction. The Adaptive Learning Environments Model classes were observed to feature constructive student interactions, encouragement of self-management, student choice, exploration, and the teacher acting as manager and consultant rather than as disciplinarian or lecturer for the whole class or small groups. In the Team-Assisted Individualization classes, students were observed to work individually on written assignments and tests and quizzes; the teacher's role was to diagnose and assist.

As a further example of how results from analyses of the patterns of similarities and differences in program features can be used to analyze the quality of the learning environment under each program, the observation data are summarized in Figures 1-A through 1-D. For illustrative purposes, the eight programs are referred to by letters in the figures. Figure 1-A shows, for example, that Program A was tightly clustered around the mean T-score of 50 and appeared to be the most typical or representative of the eight programs; that is, it was neither positively nor negatively distinct from the other programs. Program B, although slightly above the mean in other respects, was notable for scores below the mean on the variables, learning centers and materials in order. Program C was sharply above the mean in all five physical design features. Programs D, E, and G were clustered close to the mean in most respects. Program F had notably high scores on all features except classroom arrangement, on which it was two standard deviations below the mean. Program I was fairly low in all physical design features.

Relational Patterns

The results from a series of canonical correlation analyses of the data on program features and classroom processes suggest that when controlling for socioeconomic status, program features were closely associated with classroom processes, and both sets of variables predicted students' perceptions of classroom climate. Program features alone were found to predict students' perceptions of self-responsibility; classroom processes alone were found to predict students' adjusted achievement outcomes. Furthermore, results from the canonical correlations specifically linked program features and classroom processes. Eleven program features — student choice, task flexibility, teacher monitoring, peer tutoring, students seeking adult help, record keeping, classroom arrangement, task directions, learning centers, variety of materials, and clear labeling — were associated with classroom processes such as student use of exploratory materials, student work in parallel groups, teachers interacting with students on personal matters, student self-management, and student participation in presentations. It is noteworthy that the same 11 program features were found to be associated negatively with classroom processes such as students working in group-interactive settings, whole-class

Figure 1–A: PROGRAM FEATURE PROFILES:
Physical Design of the Classroom

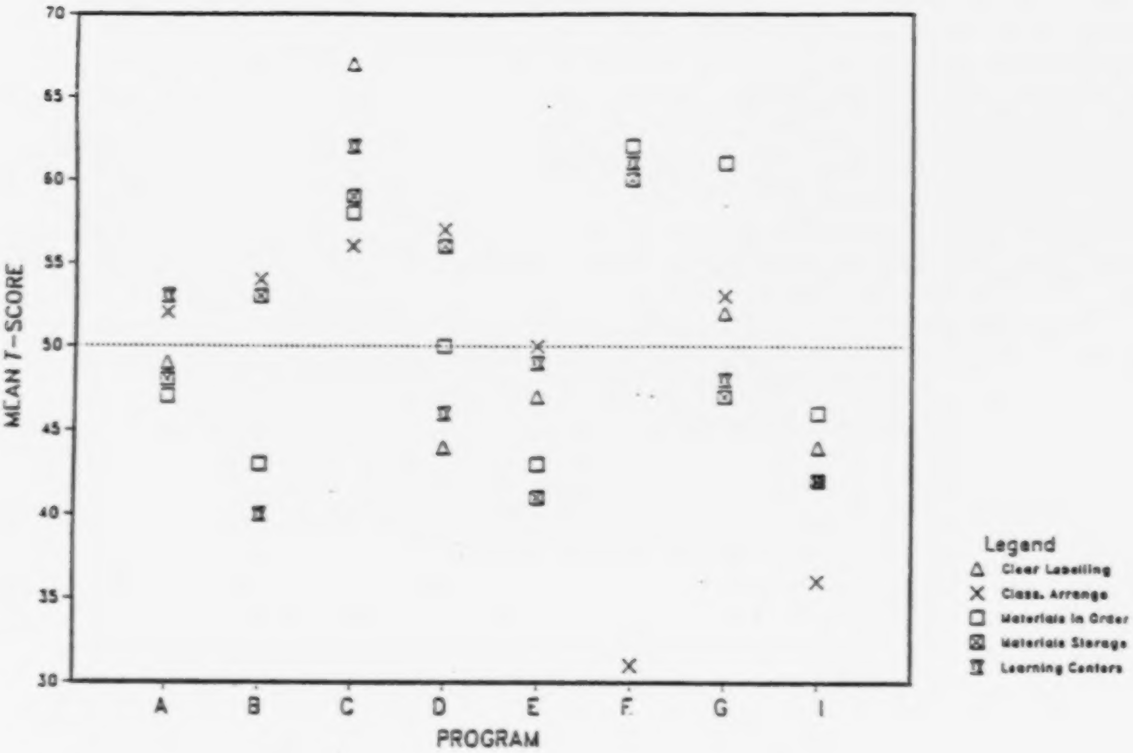


Figure 1–B: PROGRAM FEATURE PROFILES:
Curriculum Materials

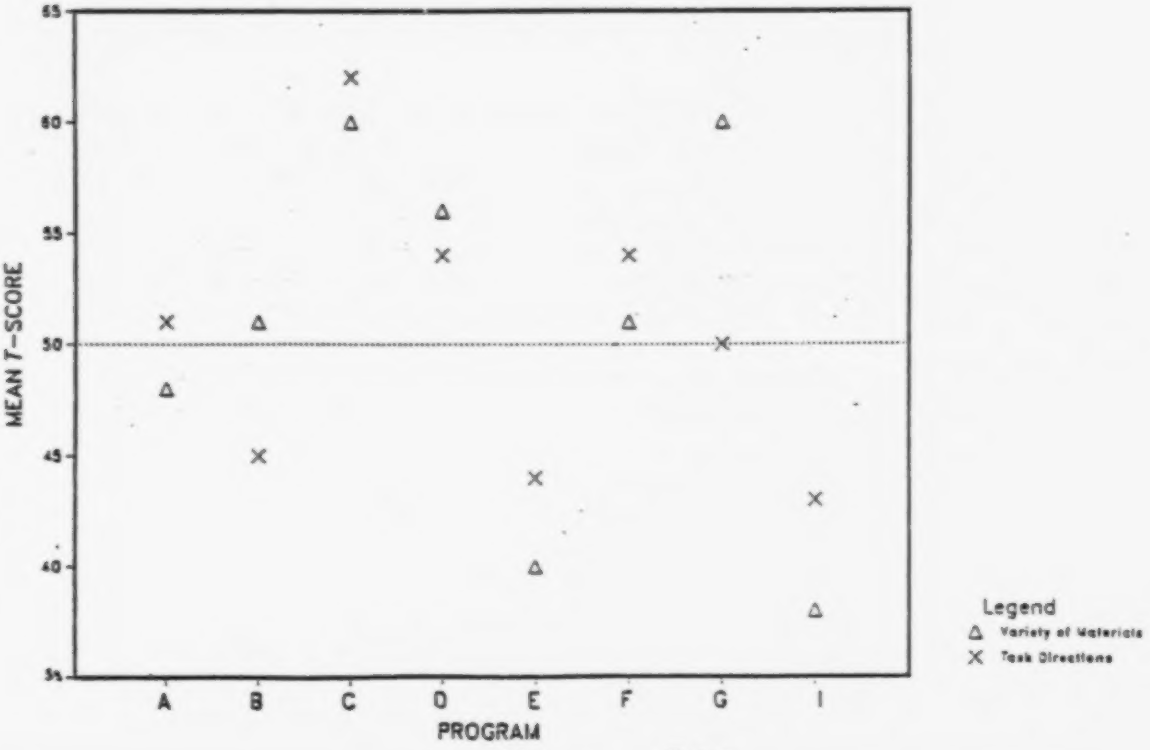


Figure 1–C: PROGRAM FEATURE PROFILES:
Provision of Adaptive Instruction (a)

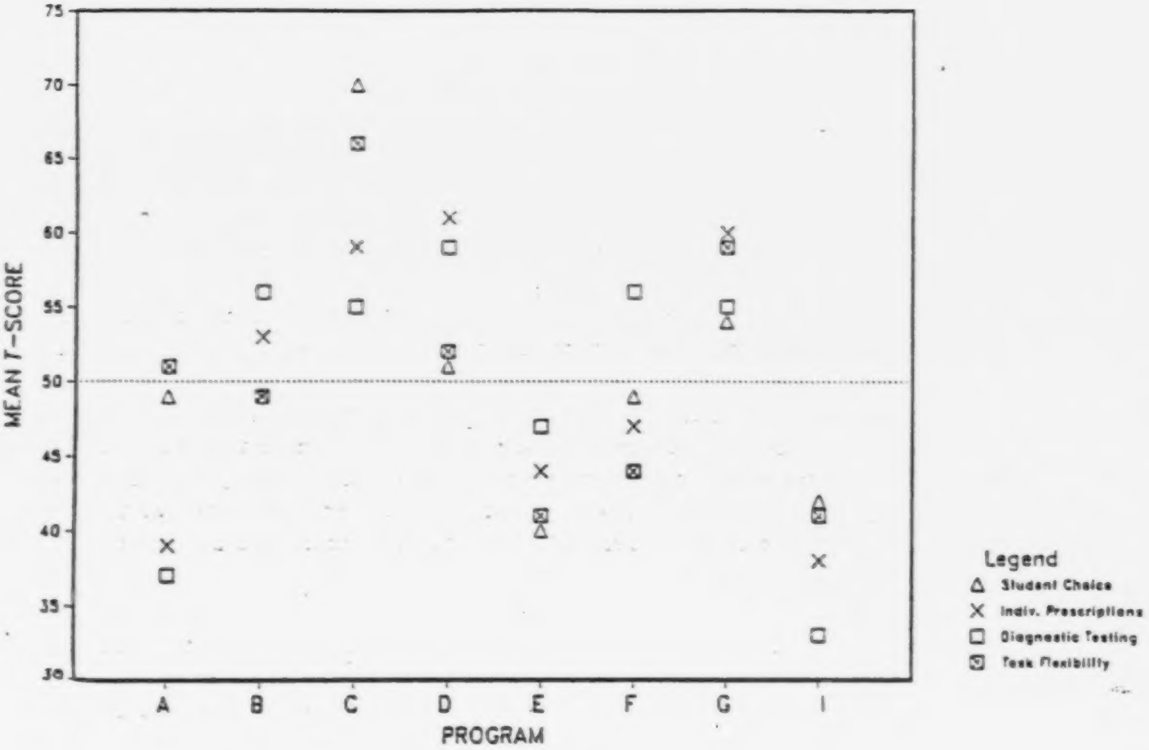
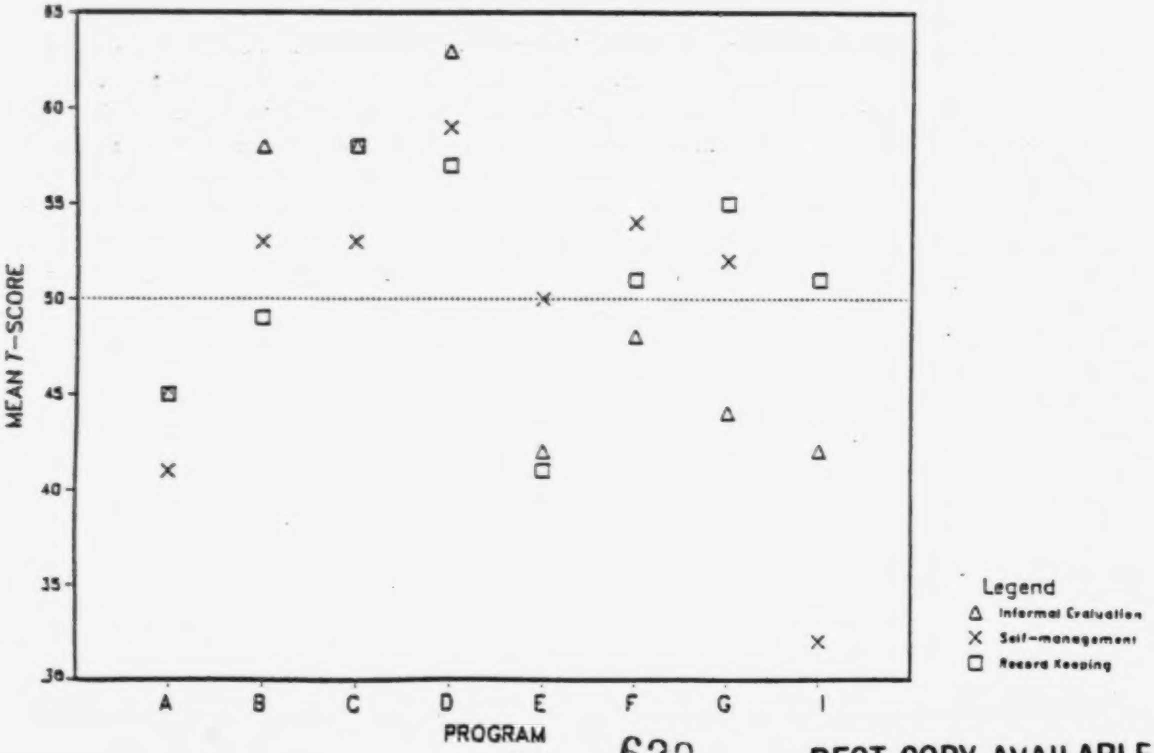


Figure 1–D: PROGRAM FEATURE PROFILES:
Provision of Adaptive Instruction (b)



BEST COPY AVAILABLE

and small-group instruction, teacher instruction, teacher questioning, and no use of materials. Also of interest are the results showing that classes which were observed to combine the program features, clear labeling, learning centers, less teacher monitoring, informal evaluation, students seeking adult help, and diagnostic testing, also were observed to be characterized by the classroom processes, constructive student interactions, student participation in presentations, less teacher explaining and cueing or prompting, and interactions with students for instructional purposes. The overall patterns of relationship between program features and classroom processes suggest the contrast between instructional features that are adaptive to student differences (e.g., one-to-one and small-group tutoring) and the traditional instructional practices that have been predominant since the turn of the century (e.g., the recitation method of questioning, teacher instruction in whole-class or small-group instructional settings). In addition, the results show that classrooms featuring the greatest use of individualized prescriptions, task flexibility, students seeking adult help, a variety of materials, and clear delineation of task-specific directions were associated most closely with high levels of student responsibility. These program features also were associated with greater student perceptions of competitiveness and friction, and lesser student perceptions of cohesiveness and satisfaction.

Findings from the analyses also show that, on average, programs with more features of adaptive instruction tend to raise achievement and student self-responsibility to levels as great as, or better than, those found for programs that feature more teacher-directed and group-paced instructional strategies. In addition, several of the programs with the primary goal of providing for student differences produce superior classroom processes that many students, parents, and educators greatly value. These include constructive student interactions, independent work, individual diagnosis and prescription, cooperative learning, student exploration, and teachers interacting with students on personal matters.

Overall, findings such as those discussed above make it possible to delineate relationships between specific program features and classroom processes and student learning outcomes. The results show that, when they are well-implemented, features such as the allocation of available class time for curriculum-related activities, a variety of instructional strategies, a variety of materials and activities, and learning tasks that are appropriate for students' learning needs and achievement levels, can produce superior classroom processes and achievement results that are not unlike those associated with ideal realizations of traditional, teacher-directed and group-paced instruction. Moreover, features such as student choice, which is suggested in the effective-teaching literature to be an ineffective feature of adaptive instruction programs, actually were found to facilitate student learning.

In light of the findings from the observation study discussed above and from the quantitative synthesis of adaptive instruction programs by Waxman et al. (1985) described earlier in this paper, and given the current push for educational excellence and basic skills acquisition, it seems critical to begin accumulating further evidence that verifies, or contrasts with, the predominant literature supporting the efficacy of group-paced, teacher-directed instruction. One of the central arguments of the extant effective-teaching literature is that programs which make allowances for individual differences, student initiative, and student self-responsibility for learning tend to be ineffective in increasing time-on-task and basic skills acquisition, while also being impractical for widespread implementation in regular classroom settings (e.g., Bennett, 1976; Brophy, 1979; Hedges, Giaconia, & Gage, 1981). Findings from the observation study by Wang and Walberg, as well as those from the Waxman et al. (1985) quantitative synthesis, are a counterpoint to this argument. In particular, the results from the Wang and Walberg study are pertinent for two important reasons. First, they illustrate that high degrees of implementation of adaptive instruction features in regular classrooms can be established and maintained in a variety of school settings. Second, they show that although different instructional approaches were in use, there seemed to be a close resemblance in observed classroom processes between the exemplary classrooms of the programs included in the study and instructionally effective classrooms as portrayed in the effective-teaching research literature. In this context, these studies represent an important step toward accumulating the kind of data base on instructional features that currently is sorely lacking. Information of this sort is critical for making informed choices from among alternative educational models and for strategic planning aimed at creating alternative futures scenarios with the goal of improving current practice.

CONCLUSION

This paper presents a case for gathering information on specific features of school learning environments that are effective in maximizing all students' chances for schooling success, particularly the chances of success for students who require greater-than-usual educational support. Examination of research and practice supports the contention that information on learning environments or conditions, combined with a broadened data base on student outcomes, can greatly enhance innovative program development, school implementation, and strategic planning. Recent studies with implications for increasing the effectiveness of schooling come from virtually all areas of research on human development and learning, as well as from investigations of effective teaching and classroom processes. This research is adding substantially to our understanding of learner competence, of how such competence is acquired, and of some key characteristics of effective schooling.

The important leadership role of NCES is implicit in the recommendations discussed in this paper for building upon the current research base. This role dictates an improvement orientation whereby data on the quality of education are used (a) "to assist educators and educational policy makers by informing their decisions and to assist the general public by describing the 'health' of American education" (U. S. Department of Education, 1985); and (b) to create alternative futures scenarios and conduct related strategic planning. Thus, NCES is called upon to greatly expand its current focus on disseminating information for the purposes of trend analysis and forecasting; it is challenged to accept responsibility for increasing the relevance of its data base to educators and policymakers as well as providing the general public with information that can be used in making informed choices from among different schooling approaches: The ultimate goal is for NCES to make available information that is most timely and relevant to informed decision making by educational planners and informed choices by the public -- the primary consumers of educational improvement.

This is a special period in American education. The pervasiveness of the sense that we must somehow improve the quality of schooling is reflected in over 30 major national reports and in the creation of over 300 state task forces on the general quality of education (Cross, 1984). Improvement efforts are under way, and many of them have considerable potential. The current wave of educational reform underscores the leadership role of NCES in pooling the resources and capabilities of other federal, state, and local information-gathering agencies. Only in this way can progress be made in systematically building data bases that go beyond the predominant focus on the socioeconomics of education and address issues related to improving the quality of educational practice.

REFERENCES

Becker, W. C., Engelmann, S., Carnine, D. W., & Rhine, W. R. (1981). Direct Instruction Model. In R. W. Rhine (Ed.), Making schools more effective: New directions from Follow Through. New York: Academic Press.

Bennett, N. (1976). Teaching styles and pupil progress. Cambridge, MA: Harvard University Press.

Bloom, B. S. (1968). Learning for mastery. Evaluation Comment, 1(2), 74-86.

Brandt, R. S. (Ed.). (1985). Success for all students. Educational Leadership (Special Issue), 43(1).

Brophy, J. (1979). Teacher behavior and its effects. Journal of Educational Psychology, 70, 733-750.

Brown, A. (1978). Knowing when, where, and how to remember. In R. Glaser (Ed.), Advances in instructional psychology (Vol. 1). Hillsdale, NJ: Erlbaum.

Cantalician Foundation, Inc. (1983). Technical assistance and alternative practices related to the problem of the overrepresentation of Black and other minority students in classes for the educable mentally retarded. Buffalo, NY: Author.

Cross, K. P., (1984). The rising tide of school reform reports. Phi Delta Kappan, 66(3), 167-172.

Doyle, W. (1979). Classroom tasks and students' abilities. In P. L. Peterson & H. J. Walberg (Eds.), Research on teaching: Concepts, findings, and implications. Berkeley, CA: McCutchan.

Gilkeson, E. C., Smithberg, L. M., Bowman, G. W., & Rhine, W. R. (1981). Bank Street Model: A developmental-interaction approach. In W. R. Rhine (Ed.), Making schools more effective: New directions from Follow Through. New York: Academic Press.

Glaser, R., & Bond, L. (1981). Testing: concepts, policy, practice, and research. American Psychologist, 36(10), 997-1000.

Good, T. L. (Ed.). (1985). Policy implications of effective schools research. The Elementary School Journal (Special Issue), 85(3).

Hedges, L. V., Giaconia, R. M., & Gage, N. L. (1981). Meta analysis of the effects of open and traditional instruction. Stanford, CA: Stanford University Program on Teaching Effectiveness.

Heller, K., Holtzman, W., & Messick, S. (Eds.). (1982). Placing children in special education: A strategy for equity. Washington, DC: National Academy of Sciences Press.

Jeter, J. (Ed.). (1980). Approaches to individualized education. Alexandria, VA: Association for Supervision and Curriculum Development.

Klausmeier, H. J. (1972). Individually Guided Education: An alternative system of elementary schooling. Harlan E. Anderson Lecture, Yale University Center for the Study of Education, New Haven, CT.

Larry P. v. Riles, Civil action N. L. - 71-2270, 343 F. Supp. 1306 (N. D. Cal., 1972).

Mayor's Commission on Special Education. (1985). Special education: A call for quality. New York: Author.

National Coalition of Advocates for Students. (1985). Barriers to excellence: Our children at risk. Boston: Author.

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. Washington, DC: U.S. Department of Education.

National School Public Relations Assocation. (1981). Good schools: What makes them work? Arlington, VA: Author.

Plisko, V. W. (Ed.). The condition of education (1984 Edition). Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Ramp, E. A., & Rhine, W. R. (1981). Behavior Analysis Model. In W. R. Rhine (Ed.), Making schools more effective: New directions from Follow Through. New York: Academic Press.

Ravitch, D. (1983). Forgetting the questions: The problem of educational reform. Independent School, 42(3), 57-62.

Reynolds, M. C., & Wang, M. C. (1983). Restructuring "special" school programs: A position paper. Policy Studies Review, 2(1), 189-212.

Reschly, D. (in press). Mildly handicapped learning characteristics: Placement and instruction. In M. C. Wang, M. C. Reynolds, & H. J. Walberg (Eds.), Handbook of special education research and practice. London: Pergamon.

Rhine, W. R. (Ed.). (1981). Making schools more effective: New directions from Follow Through. New York: Academic Press.

Segal, J. W., Chipman, S. F., & Glaser, R. (Eds.). (1985). Thinking and learning skills:

Relating instruction to research (Vol. 1). Hillsdale, NJ: Erlbaum.

Slavin, R. E. (1983). Team-Assisted Individualization: A cooperative learning solution for adaptive instruction in mathematics. Baltimore, MD: The Johns Hopkins University, Center for the Social Organization of Schools.

Talmage, H. (Ed.). (1975). Systems of individualized education. Berkeley, CA: McCutchan.

Twentieth Century Fund Task Force. (1983). Report of the Twentieth Century Fund Task Force on federal elementary and secondary education policy. New York: Author.

U.S. Department of Education. (1985). Indicators of education status and trends. Washington, DC: Author.

U.S. House of Representatives. (1983). The report from the Commission on the Financing of a Free and Appropriate Education for Special Needs Children. Washington, DC: Author.

Walberg, H. J. (1984). Improving the productivity of America's schools. Educational Leadership, 41(8), 19-30.

Wang, M. C., Birch, J. W., Anderson, K. A., & Reynolds, M. C. (1985). An empirical data base for mainstreaming. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center.

Wang, M. C., Gennari, P., & Waxman, H. C. (1985). The Adaptive Learning Environments Model: Design, implementation, and effects. In M. C. Wang & H. J. Walberg (Eds.), Adapting instruction to individual differences. Berkeley, CA: McCutchan.

Wang, M. C., & Lindvall, C. M. (1984). Individual differences and school learning environments. In E. W. Gordon (Ed.), Review of research in education (Vol. 11). Washington, DC: American Educational Research Association.

Wang, M. C., & Peverly, S. T. (in press). The role of the learner: An individual difference variable in school learning and functioning. In M. C. Wang, M. C. Reynolds, & H. J. Walberg (Eds.), Handbook of special education research and practice. London: Pergamon.

Wang, M. C., Rubenstein, J. L., & Reynolds, M. C. (1985). Clearing the road to success for students with special needs. Educational Leadership, 43(1), 62-67.

Wang, M. C., & Walberg, H. J. (Eds.). (1985). Adapting instruction to individual differences. Berkeley, CA: McCutchan.

Wang, M. C., & Walberg, H. J. (in press). An investigation of classroom processes of exemplary adaptive instruction. *Monograph Series of the International Journal of Educational Research*. London: Pergamon.

Waxman, H. C., Wang, M. C., Anderson, K. A., & Walberg, H. J. (1985). Adaptive education and student outcomes: A quantitative synthesis. *Journal of Educational Research*, *78*(4), 228-236.

Wittrock, M. C. (Ed.). (in press). *Third handbook of research on teaching*. Washington, DC: American Educational Research Association.

Ysseldyke, J., Algozzine, B., & Epps, S. (1983). A logical and empirical analysis of current practice in classifying students as handicapped. *Exceptional Children*, *50*, 160-166.

Assessing Students' Social and Communicative Competence in School

Louise Cherry Wilkinson
City University of New York, Graduate School

As part of the national effort to measure the progress of elementary and secondary students in American schools, their social and communicative performance should be assessed, and the classrooms in which they acquire knowledge should be described. I share the position held by many educators that the educational statistics collected about our children and about our schools have had limited utility and questionable validity.

Past surveys of the status of American education have not fully described, nor have they validly predicted the ways in which students learn and achieve in school. One reason for this failure is that the prior analyses have not included descriptions of classrooms. Another reason is that they have not included measures of important aspects of school-age children's social and communicative knowledge. Social knowledge includes the ability to perform functionally appropriate interpersonal behaviors, while communicative knowledge includes the ability to use verbal and vocal expression to communicate. Achieving a sense of autonomy and responsibility, communicating effectively with others, cooperating with others and solving interpersonal problems, are all within these domains of knowledge. They change dramatically during the school-age years and are heavily influenced by students' experiences in classrooms. Social and communicative competence are important achievements that the educational system can either foster or discourage. They should, therefore, be included in descriptions of American education.

In this paper, I comment on the present state of statistical descriptions of American education. Then I discuss the need for new measures that assess the social and communicative competence of students and the climate of their classrooms. Finally, I offer some suggestions for more comprehensive assessments.

Descriptions of Students' Achievement and Their Classrooms

Over the years, large amounts of information have been collected on student characteristics and student outcomes as indicators of the state of American education. It is appropriate and logical that these descriptions focus on students' academic achievement in specific curricular areas, such as arithmatic or reading, and basic verbal, spatial, and mathematical aptitudes. Past reports by NCES are replete with descriptions such as the following: "SAT Scores for Students 1963-1984," "Eighth grade Mathematics Achievement," "National Assessment of Educational Progress in Reading for Ages 9, 13, 17, by Selected Characteristics of Participants in the United States, 1974-1975, and 1979-1980." These reports are deficient in two areas. (1) They neglect social and communicative competence. (2) They contain virtually no information on the classrooms in which students are taught. There is modest information about students with limited English proficiency and some information on foreign-language enrollments, such as the following: "The Percent of 5 to 17-year-olds that Speak a Language Other Than English at Home," "Estimated Language Minority Status and English Language Proficiency of Population 5 to 14-year-olds, 1976 to 1982," and "The Mean Number of Carnegie Units Earned by Selected Subject Matter Areas (Foreign Languages)."

Expanding Descriptions of Students' Achievement and Classrooms

One reason for measuring students' social and communicative competence in classrooms is that this knowledge mediates both the teaching and learning of academic subjects. Social and communicative competence limits the extent to which students can benefit from the instruction in academic subjects such as reading or mathematics. For example, recent research in sociolinguistic studies of classrooms provides evidence that for some students, the ways of communicating at school differ from those at home (Wilkinson, 1982). Discontinuities between the classroom and other learning environments, such as the home, may cause special problems for some students that can interfere with their overall achievement and adjustment in school. Some culturally diverse students do not know, nor are they taught the "rules of the game"--the standard ways of communicating in classrooms, such as how, when, and where to ask for and provide information to teachers and other students. The effects of students' lack of knowledge about classrooms are not limited to the obvious problems that these students face in their failure to communicate adequately. In addition to such immediate problems, if some children do not understand the classroom with its unique communicative demands, then they learn little from the instruction they receive. Inadequate learning of how to communicate in classrooms has unintended effects that include lack of understanding and inability to function in classrooms. This problem is demonstrated, for example, by students misunderstandings of the curriculum and by teachers' lowered expectations that can then result in their differential treatment of students. Furthermore, accurate assessment of the student's achievement is unlikely, since access to their knowledge is predicated upon good communication. Perhaps the educational failure of some students, as documented by educational statistics, is caused in part by differences in communication between students and teachers who come from different cultural backgrounds.

My studies of students and classrooms provide an example of the interrelationship among academic, social, and communicative achievements. During the past several years, my colleagues and I have been especially interested in how students provide information to one and other, and how they use requests and responses in instructional groups in which the teacher is often absent. Making requests, such as asking for information, and receiving adequate responses, such as correct answers, are central to teaching and learning in classrooms. Children often use requests to provide information for each other as well as to guide their behavior. Requests are common in classrooms, accounting for about two-thirds of all of the teachers' speech, and more than one-half of students' speech.

We conducted research on students' in small instructional groups and proposed a model that describes students' use of requests and responses in these groups. The central idea of the model is the concept of the effective speaker, who is successful in communicating with others. In the case of requests, an effective speaker is defined as one who receives appropriate responses to requests. Our model identifies several characteristics of requests that are associated with obtaining appropriate responses for elementary school children; these characteristics include requests that are expressed clearly and directly, that have to do with the academic task at hand, that are understood by other students as sincere, and that are revised if they do not initially result in obtaining an appropriate responses. The results of four studies of elementary school students from first through third grades that we conducted during the past several years provide support for our model (Wilkinson, 1982, 1983). Requests that conform to the model are most likely to result in students' obtaining appropriate responses during their reading and mathematics in small, instructional groups. One of the most provocative findings from the research is that students who were effective in obtaining appropriate responses not only produced requests that conformed to our model but were also the highest achievers in reading and mathematics as measured by standardized tests. Our work and that of other researchers suggests that the communicative climate of the classroom may help to maintain and encourage differences among students in their mathematics and reading achievement. Initial differences among students in their academic achievement and communicative knowledge may contribute to what they actually learn in classrooms and thus affect their performance on standardized tests.

Assessing Students' Social and Communicative Competence in Classrooms

We should develop and administer assessments of students' social and communicative competence in classrooms. A set of educational statistics that purports to describe the state of American education should include measurements of the social and communicative competence of school-age children. I will suggest a way of developing adequate indicators that accurately reflect students'competence in these areas.

The task of measuring social and communicative competence of school-age children in school is difficult, because there is no standardized, group-administered test that accurately and comprehensively assesses these competencies. To the contrary, research suggests that measurement of these competencies is more accurately reflected by direct observation of students' actual behavior in classrooms than by test scores.

Any attempt to measure social and communicative competence should be guided by the current state of scientific knowledge. The design of assessments should be guided by contemporary theory in child development and learning. For example, one could assess children's ability to solve social problems (e.g. Chandler, 1973), to communicate effectively (e.g. Krause & Glucksberg, 1969), to reason about complex moral issues (e.g. Kohlberg, 1976), to know about the social relationships in classrooms (e.g. Cohen, 1984), and to understand self-concept (e.g. Harter, 1984).

It is essential, however, that assessments also include samples of students' natural behavior as it occurs in classroom. For example, we have constructed profiles for each student based on his/her communicative effectiveness in one classroom situation: the small, instructional group. In our research (Wilkinson, 1982, 1983) the profile includes the following: (1) the conditions under which the student's requests are successful in obtaining information, (2) the variety of ways in which the student tries to obtain information, and (3) the student's response if the information is not obtained. After collecting such profiles, we must consider the norms against which the profiles are compared. Importantly, we must be sensitive to differences that are primarily due to the individual, versus the culture (e.g. dialect) or the situation (e.g. a classroom situation when the teacher is absent).

The design of future assessments of students' achievement in classrooms should include both longitudinal and cross-sectional designs, so that trends in students' social, communicative, and academic achievement can be noted. If there is a trend of note (e.g. a decline in third-grade students' reading scores), it may be worthwhile to collect more information.

Assessments should be administered early in students' school years, for two reasons. (1) Educators and policy makers need information on the state of young children to guide decisions about early childhood education. (2) Contemporary research in child development and learning has established that subsequent achievement in all domains of knowledge is heavily influenced by students' early learning and experiences in and outside of classrooms. It would be helpful to understand the root of a trend to guide subsequent decision making.

Conclusion

In sum, future reports on American education should include the following. (1) They should address students' social and communicative competence in addition to their achievement in specific curricular areas, such as reading or mathematics. (2) They should include direct observation of students' naturally occurring behavior in a variety of classroom situations. (3) They should carefully take into account differences among students that are due to cultural and situational factors, and true individual differences among students.

NCES should play a strong role in developing and administering a new set of assessment tools that more adequately reflect the state of American education at any point in time. Finally, any effort to chronicle American education that extends into the next century must include a national committment to support basic research into children's learning and development in classrooms. This is a necessary investment if we are to have the needed tools in the future to carry on with this important endeavor.

References

Chandler, M. (1973) Egocentrism and antisocial behavior: The assessment and training of social perspective-taking skills. Developmental Psychology, 9, (3), 326-332.

Cohen, E. (1984) Talking and working together: Status, interaction, and learning. In P. Peterson, L. C. Wilkinson, & M. Hallinan (Eds.) The social context of instruction, Orlando, FL: Academic Press.

Harter, S.(1984) Developmental perspectives on the self-system. In M. Hetherington (Ed.) Handbook of Child Psychology, Volume IV Socialization, Personality, and Social Development, 4th Edition, New York, Wiley.

Kohlberg, L. (1976) Moral stages and moralization: The cognitive-developmental approach. In T. Lickona (Ed.) Moral development and behavior. New York: Holt, Rinehart, & Winston.

Wilkinson, L. Cherry (1982) Communicating in the classroom. New York: Academic Press.

Wilkinson, L. Cherry & Calculator, S. (1982) Requests and responses in peer-directed reading ability groups. American Educational Research Journal, 19 (1), 107-120.

Wilkinson, L. Cherry & Spinelli, F. (1983) Using requests effectively in peer-instructional groups. American Educational Research Journal, 20 (4), 479-502.

American Association for
Counseling and Development
5999 Stevenson Avenue, Alexandria, Virginia 22304 703/823-9800

Serving the counseling,
guidance and human
development professions
since 1952.

June 18, 1985

Mr. Emerson J. Elliott, Administrator
National Center for Education Statistics
Brown Building
1200 19th Street, N.W.
Washington, D.C. 20208

Dear Mr. Elliott:

The American Association for Counseling and Development (formerly the American Personnel and Guidance Association) is pleased to participate in the redesign of the National Center for Education Statistics elementary and secondary education data program. There are some important data collection/research issues that should be addressed in the future to provide invaluable information about learning and the rules and responsibilities of school personnel.

Rather than provide lengthy papers on these issues, we have chosen to raise them in this letter, with some suggestions when possible. We would be pleased to discuss them in more detail with you and to devote staff and reader time to your redesign effort.

Demographic Data

In the past, there have been a number of problems in collecting data on noninstructional personnel. For example, elementary school counselors often serve more than one school, making it possible, perhaps even likely, that the same person is counted more than once. This problem is true of nurses, social workers and others as well. It is essential that the number of noninstructional personnel be reported on a district basis in terms of full-time equivalent units. It would also be valuable to know what the ratio of students to these different personnel is in each district. Because the quality of services provided to students rests upon the caliber of instruction as well as the types and extent of student needs that are met, it is essential to assess accurately the number and types of noninstructional personnel who provide services to students. AACD worked collaboratively with NCES in the preparation of Counselors in Local Education Agencies, Fall 1979 and Trends Since 1970 (NCES 82-122b). We welcome the opportunity to participate in the updating of this research report.

Time and Task Analysis

Recent studies on how teachers and students spend their classroom time have not only been instructive in understanding what is happening in the classroom but have led educators to reassess how they spend their time. Similar studies for noninstructional personnel would be invaluable in assessing the types of services provided and the actual needs of students and faculty. Educators at all levels could use this information to:

1. Gain a better understanding of the roles of noninstructional personnel, including (but not limited to) school counselors, psychologists, social workers, nurses, speech and hearing specialists, librarians and media specialists, administrators and supervisors;

2. Assess student needs and the degree to which they are being met by noninstructional personnel;

3. Review and, if necessary, reformulate the roles and responsibilities of noninstructional personnel; and

4. Develop a comprehensive approach based on realities, not perceptions, of the appropriate tasks of all school personnel.

Note: AACD is extremely interested in and willing to assist in any time and task studies directed to the study of the school counselor's role in elementary and secondary education.

Demographic Data on Educational Personnel

Information on education personnel, their ages, years of experiences, training and average length of stay is essential if we are to predict future personnel needs in a more systematic manner. These demographic data need to be correlated with expected student enrollment, based on birth and enrollment statistics, to project future personnel needs. These projections should address the needs at least 5-10 years in advance to encourage/discourage youth and adults from pursuing training as educators. The projected teacher shortage illustrates the importance of such planning. Systematic, accurate information on current personnel and future needs is vital for all education personnel if we are to lessen or prevent personnel shortages or surpluses in the future.

Note: AACD can assist NCES in assessing the status of student enrollment in elementary and secondary counselor education programs. The association is also committed to gaining better supply and demand information for counseling and pupil services personnel.

Career Development

Because the goal of education is not only to produce an educated citizenry but also to prepare future generations to enter the work force, it is essential that we assess students' career development needs, current activities of schools in fostering career development and what might be done to improve this aspect of school preparation. Such an assessment would require an indepth study, but that fact should not be a deterrent to it. The future of our children, our economy and our country is intricately tied to successful employment of future generations.

Need for NCES Advisory Council

We strongly recommend that NCES institute an advisory council composed of educators from diverse instructional and noninstructional backgrounds, as well as experts in assessment, parents and students to provide advice on future survey efforts. This advisory council should be a formalized body, meeting to offer advice and direction to NCES. Diverse instructional and noninstructional personnel should be involved, regardless of the survey's focus, because few, if any, research areas are restricted to only one type of personnel. This advisory council should provide direction and advice and, in some cases, be involved in responding to the survey design, draft items and the method(s) for survey dissemination. Such involvement can help identify problems of overlap, unclear questions, misinterpretation of terminology, gaps and other issues that ultimately reduce the validity and reliability of the data collected.

Note: AACD is willing to identify staff, leader and professional representatives to participate in such an NCES advisory council.

Assessing School Counseling

To support the belief that the Department and AACD should work in conjunction with each other, we have enclosed a list of research questions for assessing school counseling that we submitted in the fall of 1984 to the National Institute of Education as they formulated their priorities for the national center and laboratory competition. While these questions take a more indepth look at research issues related to school counseling, we think you might find them of value as you address specific questions related to school counseling.

AACD stands ready to assist NCES as the center studies and develops ideas for collecting elementary and secondary education data. We can identify professionals who have the expertise you require as you look broadly at the education services we are currently providing and those we should be providing in the future. We welcome the opportunity to provide additional information or to answer any questions about our suggestions. We wish to be involved in future efforts.

Sincerely yours,

Patrick J McDonough

Patrick J. McDonough, Ed.D., NCC
Executive Director

FEB:LH
Enclosure

American Association for
Counseling and Development
5999 Stevenson Avenue, Alexandria, Virginia 22304 703/823-9800

Serving the counseling,
guidance and human
development professions
since 1952.

RESEARCH QUESTIONS FOR ASSESSING SCHOOL COUNSELING

1a. A comprehensive accountability study of existing guidance programs is needed. It should:

  (a) identify real and implied programmatic goals, measure how they are being accomplished and how successful they are;
  (b) assess how counselors spend their time and analyze the costs and benefits of this time use;
  (c) survey opinions of various consumers about the quality of counseling services being delivered and their perceived needs for counseling services; and
  (d) assess what school counselors do most effectively.

2. What is the impact of counselors' primary prevention efforts? Examples of such programs include career education, relationship enhancement, parent and teacher effectiveness training, peer counseling, assertiveness training, coping skills training and problem solving training.

3. What is known about effective decision-making? Can children/adolescents be taught to be effective decisionmakers? Does such training have an important impact on their lives? What is the most effective way to teach these skills?

4. Can children/adolescents be taught effective problem-solving skills? What are the most effective ways to do so? Are they able to apply these skills to their own lives? Does such training have an important impact on their lives?

5. What is the relationship of temperament style and characteristics to various approaches of counseling, guidance and/or learning? In other words, what kinds of interventions are most successful with what kinds of personalities?

6. What are some of the problems with cross-cultural counseling and how can they be overcome? What counseling techniques work best with different types of students (e.g., minorities, disadvantaged, etc.)

7. Can counselors make a significant impact on children's/adolescents' achievement through: (a) individual counseling, (b) groups approaches, or (c) consultation with teachers and parents? Which approach is most effective in which situations?

American Association for Counseling and Development is Committed to Equal Opportunity.

8. How does the learning environment affect learning and can counselors change the learning environment to be more responsive to students' needs?

9. What is the role of the counselor in identifying children who are at-risk for personal, educational, or career problems? How can teachers and counselors work more effectively with children who have problems?

10. What are the most effective approaches for the counselor in assisting students and their parents with career exploration and selection?

11. What interventions are most effective in working with the following:

    - learning disabled
    - emotionally disturbed/deprived
    - physically handicapped
    - children of divorce
    - children in step-families
    - economically disadvantaged
    - substance abuse
    - child abuse
    - premature or delayed development
    - chronic truancy
    - underachievement
    - social immaturity

For further information, please contact Dr. Frank Burtnett or Dr. Sharon Alexander at AACD.

Comments for the
National Center for Education Statistics
Redesign of the Elementary and Secondary Education Data Program

June 14, 1985

From the

American Association of Colleges for Teacher Education

Prepared by

Elizabeth A. Ashburn, AACTE Director of Research and Information Services

Edward Ducharme, Chair, Organizational Counseling and Foundations Studies, University of Vermont; Member, AACTE Task Force on Research and Information

Kenneth Howey, Associate Dean and Professor, College of Education, University of Minnesota; Member, AACTE Task Force on Research and Information

David G. Imig, AACTE Executive Director

David C. Smith, Dean, College of Education, University of Florida; AACTE Immediate Past President and Member, AACTE Task Force on Research and Information

Sam J. Yarger, Dean, School of Education, University of Wisconsin-Milwaukee; Chair, AACTE Task Force on Research and Information

Nancy L. Zimpher, Professor, The Ohio State University; Member, AACTE Task Force on Research and Information

AACTE gathers data about higher education-based teacher education; consequently AACTE does not consider itself expert in the area of data collection concerning elementary or secondary education. Our membership has been involved, however, with a variety of data collection vehicles sponsored by the National Center for Education Statistics. On the basis of this involvement and the Association's data-collection efforts in another sector, the following comments and observations are offered.

General Areas of Data Needs

In examining the documents distributed by NCES in this call for comment, and taking into consideration some long-term data needs of the teacher education community, we have identified four areas which should have a high priority for NCES:

- Teacher supply and demand. With increasing competition for scarce resources at both the preservice teacher education and inservice levels, it becomes imperative to have accurate current information on and future projections of teacher supply and demand. Such scarce resources need to be distributed so that teacher education programs can be responsive to the school personnel needs of local districts. Information on the teacher reserve pool (its size, mobility, and interest in returning to teaching), the retirement picture for the current workforce, accurate attrition figures, and "lateral entry" forecasts are needed both short-term and long-term.

- Beginning teacher induction programs and inservice education. Data about programs in these areas are critical because they have implications for future program development. Estimates as to growth (or lack thereof) of inservice and beginning teacher induction programs for teachers will allow teacher educators to prepare intelligently to assist school districts with the continuing education of teachers.

- Data about teacher testing programs. To a large degree, the credibility of teacher preparation institutions is on the line with teacher tests. Despite the perceived inadequacies of the current tests, the quality of schools, colleges and departments of education is likely to be judged based on a variety of competency tests. If support for rigorous and demanding tests can be demonstrated, a "professional school mentality" may start to develop. In other professions, e.g., accounting and law, it is normal for 30, 40, or even 50% of the applicants to not pass the test on the first attempt; despite these passage rates, the training institutions are rarely blamed. The relationship in those circumstances is between the testing agency and the prospective professional. Typically, the training institution will offer programs to help students who are having trouble passing those tests prepare to do better on the next try. A similar attitude with respect to the teaching profession is necessary and desirable. The more information that can be obtained on teacher testing programs, the better teacher education programs will be able to prepare to meet the needs of teachers in this regard.

- Continuity of data gathering. We emphasize strongly that most of our recommendations for data-gathering will have little consequence if data are not gathered on a continuing and systematic basis over long periods of time. Frequently, what is needed is trend data, not individual data for a given year, since of overwhelming concern to the education community is continuing quality and meeting future education needs with present teacher preparation.

Sources of Data

NCES gathers and organizes data in three principal ways, according to Attachment B of the information provided in your request to us. First, a variety of data-gathering arrangements are in place in state education agencies. Second, the Center sponsors voluntary response to sample surveys conducted by mail. Finally, NCES gathers data collected by other federal agencies. We have two general comments about these approaches to data collection.

Voluntary Mail Surveys. Our observation of and experience with data collection in schools of education causes some concern regarding voluntary response to sample surveys by mail. We assume, for example, that some, if not all, of the data concerning teacher supply are gathered from the complex questionnaires that are sent to selected registrars in colleges and universities around the country. Typically, these are passed on to deans of education for response. Unfortunately, many of the items requesting specific information require a response that is often too detailed and too complex for the respondent. Consequently, the questionnaires may be discarded, or more significantly, estimates may be fabricated for the purpose of appearing to be in compliance with the request. Therefore, we conclude that, given our experience, much of the data are suspect.

We recommend that NCES develop a data-gathering strategy that brings the Center into closer and more intimate contact with the potential respondent through reliance on professional associations. Recognizing that this strategy can be very costly, it would be more acceptable to allow a higher margin of error than to leave questions of accuracy unanswered. Within the constraint of scarce resources, the Center should focus on the selection of a smaller sample and take the steps necessary to enlist institutional support and involvement with the data-gathering strategy. More than likely, this would require personal contact by either a contractor or Center staff, but we believe that such contact is necessary in order to ensure the necessary respondent involvement. Even though the error margin might be larger than the ideal, the representativeness of the data is likely to be more powerful.

Collaboration. The document "Indicators of Education Status and Trends" (January 1985) lists the variety of other federal agencies and departments with which the Center works in the development of data-gathering strategies. Such efforts are to be applauded, since the richness of the information is undoubtedly enhanced by it. There was no mention, however, about collaboration with and use of data and information from the broad variety of non-federal sources. In the area of elementary and secondary education, for example, such organizations as the National Education Association, the American Federation of Teachers, the Council of Chief State School Officers, the National Association of Secondary School Principals, and the American Association of School Administrators, as well as others, are continually gathering information about their enterprises. We recommend that the Center initiate long-term collaborative relationships with these groups that would allow for an outlet for the important data which they collect, and also ultimately an increasing standardization of data that are collected by them. With the professional expertise and the broad-based access to data needs that NCES possesses, it could be helpful in aiding organizations to focus their limited data-gathering strategies. This influence would lead to a richer national data pool that would provide practitioners with more, and more accurate, information, and would also help reduce the number of problems that are encountered when one set of data appears to be contradicting another set. We offer no master plan concerning how these long-term collaborative relationships might evolve, but we remain convinced that such efforts would be worthwhile.

Obviously, the same idea concerning collaborative relationships can and should apply to the gathering of data beyond the elementary and secondary education program. These types of relationships could also be developed in the areas of post-secondary education, vocational education, and education of a variety of special programs and populations that are of interest to NCES and to America's educators.

Current NCES Data-Gathering Efforts

Three aspects of the current elementary and secondary program of particular concern to teacher education are discussed below in detail. With respect to the State Aggregate Fiscal Report and the Public School Survey, we believe that additional useful information could be obtained via these instruments than is currently being collected. With respect to the Survey of Teacher Demand and Shortage, a revised data-collection methodology is suggested.

State Aggregate Fiscal Report. NCES is committed to gathering state aggregate fiscal data concerning current expenditures by major function. Suggested examples are instruction, support services, and non-instructional services; it is unclear whether there are more categories than those. A category system should be added that allows data to be gathered concerning the state aggregate commitment to teacher education. This could include expenditures for inservice education, beginning teacher induction programs, and support for intern teachers as well as preservice teacher education. There is precedent for public school monies being used to support all of these classifications.

Public School Survey. It is important for the education community to know about the nature of teacher training beyond preservice teacher education. Do teachers take college courses as the primary strategy for inservice education? Do they enroll primarily in district-sponsored inservice education? Do they enroll primarily in inservice provided by their teacher organization or by other professional organizations? Are there more informal types of inservice training that teachers use? In addition, perhaps through the data provided by the districts, it would be helpful to find out the amount of money devoted by districts to inservice education as a proportion of their total budget. It is important to know the types of structures that school districts organize in order to provide additional training for teachers. With the current thrust toward helping beginning teachers adjust to the classroom, it is important to start gathering data about the fiscal support for new teachers (induction programs). Finally, it would be very helpful if the Center could develop strategies for gathering data concerning the nature and frequency of relationships that local school districts have with institutions of higher education, focusing on inservice education.

Survey of Teacher Demand and Shortage. With regard to the strategy for obtaining information on teacher demand and shortage, we reiterate the invocation for the more intensive sampling notion we presented earlier. In the current plan, 3540 educational institutions, 2540 LEAs, and 1000 private schools are asked to provide information on the number of budgeted teaching positions, number of vancancies, etc. We have two concerns about this strategy. First, when over 7000 different people are asked to provide this kind of information, the return rate must be questioned, i.e., how many people do, in fact, respond to these surveys? Second, the completeness and accuracy of the information provided is of concern. When questionnaires come across a respondent's desk, typically from sources toward which no affiliation is felt, the tendency is to complete them as quickly as possible, often with little regard for the thoroughness and accuracy of the information provided. How does the Center ensure these data are representative? Requesting the same information from a much smaller sample via personal contact, such as a phone interview, would allow surveyors to quiz people and make judgments as to how well prepared the respondent is to deal with the question.

Other Data-Gathering Efforts

Data needs that do not appear to be adequately met by the current elementary and secondary program are described below. Teacher supply, teacher competency, and teacher career patterns are all areas where NCES could be an invaluable source of detailed data.

Teacher Supply. While no reference was found in these materials to teacher supply, and while it may be that the "supply" focus is part of the postsecondary program, it is important to mention it here. Currently, teacher supply estimates are typically made from data provided by institutions of higher education. More accurate sources of teacher supply data are state education departments: the number of teachers certified and/or licensed, the number of emergency or provisional certificates granted, the number of applications that were not granted for one reason or another, and the number of requests from school superintendents for special consideration in employing education professionals. State education department licensure figures, gathered over time, will provide trend lines on teacher supply that are much more usable than the data provided by IHEs. The reason for this is that many teacher education program graduates have no immediate goal of becoming classroom teachers, i.e., a remarkably constant number go on to graduate school, choose to stay out of the job market for reasons of marriage or family, or find alternate employment that is more appealing to them at that moment than teaching. Another confounding phenomenon is that the number of education degrees awarded does not equal the number of students newly certified to teach. Secondary education teacher candidates, for example, may have a degree in their subject area major rather than in education, and prior graduates can return for postgraduate work to obtain a teaching certificate. Thus, the number of undergraduate degrees awarded from a school, college, or department of education gives only that--an estimate of the number of people who have undergraduate education degrees; this number may be quite different, across states and across institutions, from the number of those who are actually available to teach. We recommend that NCES take an active role in collecting data about teacher supply at the state level.

Teacher Testing. NCES has a significant role to play in gathering data about the rapidly growing teacher testing movement in American education. It is important for the education community to know which tests are used and the frequency of their use.

Although not necessarily NCES's responsibility, there should be data provided concerning the validity and reliability of these tests. To the extent possible, aggregate scores by region, type of teacher, etc., should be made available. Data are needed concerning the relationship of teacher testing scores to job placement. It is also important to discover the level of discrimination that tests promote, i.e., do all teachers who take the test end up passing it or are some actually barred entrance to the profession? We recommend that NCES develop data-gathering strategies to address questions of the impact of teacher testing, in such areas as the competence of beginning and re-entry teachers, recruitment and retention of minority teachers, and the overall quality of education. Although some of these questions may go beyond the mission of NCES, we believe it is important that the questions be considered as an important context for development of longitudinal data collection efforts.

Certification and Licensure. More information is needed about certification and licensure requirements in the states. It is very difficult, at the current time, to know whether a license or certificate in one state has any relationship to that offered in another state. Reciprocity is decreasing, partly because of differences in testing programs among the various states and partly because of growing skepticism about program approval and program quality. More intensive analyses are needed of what stands behind certification and licensure requirements both across and within the various states.

We recommend that NCES play a significant role in gathering data to increase our understanding of the process of certification and licensure, a process which greatly affects the country's supply of teachers. Cooperation with the National Association of State Directors of Teacher Education and Certification could produce a ready supply of such data.

Beginning Teachers. Specific information from first- and perhaps second-year teachers over a ten-year period would be very useful for planning purposes. Knowledge about their route of entry--traditional preservice program, alternate program, lateral entry, and the type of certificate held--would be enormously helpful to policy makers. The type of support available to them as they began their teaching careers and information concerning experiences encountered in searching for a first teaching job are specific questions of concern: How many districts did they apply to? How many interviews did they receive? How many job offers did they receive? How far away from their home or college did they have to go to find a position? How far away from their first choice did they have to go? The recent spate of literature concerning the importance of the first two years of teaching to a teacher's career adds validity to these questions.

Teacher Development. By virtue of the current line of thought that teachers go through a variety of developmental phases, we recommend that the National Center for Education Statistics gather demographic, inservice education, and other data from teachers within the framework of their years of service. Such longitudinal data would help to answer important questions about degree of teacher retention, reasons for leaving teaching, reasons for remaining in teaching, development of professional competence, and impact of state and federal education initiatives in this area.

Career Ladders/Lattices. Information about career options within the teaching ranks for teachers is important, given the strong movement in elementary and secondary school systems to provide more variety and reward in teachers' work. Typically, these are referred to as career ladders or career lattices. The education community needs to know what is being done across states and what is the impact of these programs on teacher satisfaction, teacher retention, and the development of teacher competence.

Teacher Retention. The question of teacher retention has awakened considerable interest over the last five years. Although it may be possible, in a post hoc manner, to analyze some of the data gathered by NCES on this topic, it is crucial to insure that we have accurate answers to such questions as: How long do teachers stay in the teaching ranks? What are the reasons that prompt them to leave? What proportion of teachers opt or will opt for early retirement programs? What number leave and return at later stages of their lives? We recommend that NCES make this a focus for the data collection program.

Early Retirement. Another specific category of teacher demand data that would be useful concerns the prevalence of early retirement systems in schools and estimates of what proportion of teachers are taking advantage of early retirement programs. A retirement age of 65 can no longer be assumed; many states and school districts are providing incentives for people to retire early and the large number of "baby boom" era teachers are reaching an early retirement age. There would be significant benefits from estimates of future vacancies as well as from the data that the Center obtains concerning real current vacancies. This kind of information would enhance the ability of teacher educators to make long-term program development decisions.

Comment on "The Sorry State of Education Statistics" by Cooke, Ginsburg, and Smith

Cooke, Ginsburg and Smith state that education statistics as collected and published in the U.S. today are inadequate, inconsistent, incomparable, and sometimes just plain wrong. They advocate a set of "indicators" which would standardize definitions, collection parameters, and interpretations across state lines. While this is a neat theoretical solution, there are serious problems when it comes to application.

It appears that the authors are highlighting a problem that is not necessarily related to "bad" data collection, but to uncoordinated and non-structured data collection. Frequently, those who wish to make points using data from a variety of sources are not sufficiently careful to document the shortcomings of the information and, just as frequently, they misinterpret the data by being insufficiently specific about the definitions. In one example, the New Jersey and Virginia vocational education enrollment problem cited in the paper is obviously a case where the count is of headcounts of students in classes, not number of people enrolled in vocational education. If that is made clear, the data make sense, though they might not be helpful. In another example, the problem presented by the authors in understanding dropout rates appears to be a situation where the term "dropout" was probably defined differently by the census gatherers and by the school people. The authors point out reasons why people might intentionally falsify data, and that might be true, but the differences are more likely related to lack of definitional consistency.

Cooke, Ginsburg and Smith propose common definitions of indicators across states. We would agree with that, but only to the extent that that would allow either a state or NCES to have a standard by which to compare state data. Thus, when a state chooses to gather data in a different way, or chooses to gather very limited amounts of data, the statement could be made, "Their data do not fit these indicators, thus we cannot use them." In that way Secretary Bell's chart would have omitted the states of Wisconsin, Iowa, and perhaps others, because so few children took the SAT or the ACT in those states.

The paper points out that we're a very independent and autonomous country and we gather data in a variety of different ways. What the authors seem to be asking for is some system that will standardize the way terms are defined and data are collected. AACTE would submit that that is close to the collaborative relationship we have suggested. It would be unrealistic, however, as well as inappropriate, to suggest that everyone collect data in a form determined by NCES. Rather, it is likely that there will have to be careful separation of data with definitions different from those which the Center can certify as being standard and comparable. It would also be helpful for the Center to offer a data-gathering methodology critique service where they will read over any state's (or other entity's) plan for gathering information and offer advice about improvement. This strategy would be even more helpful if the Center could provide resources to help states formulate better data-gathering methods.

Summary

The following statements summarize our recommendations for the NCES ten-year plan for data-gathering from the elementary and secondary education sectors:

1. The priority data needs of the teacher education community are for information in four areas: (1) teacher supply and demand; (2) beginning teacher induction programs and inservice education; (3) teacher testing; and (4) continuity of data-gathering.

2. Collaborative relationships need to be established with professional associations and other groups with knowledge about the population being studied, and these relationships should be ongoing.

3. Current NCES data-gathering efforts should be broadened to include the areas of financial commitment to teacher education and the nature of inservice education.

4. The research design of very large samples should be modified to insure completeness and accuracy of responses.

5. Data should be collected on a continuing and systematic basis over long periods of time.

6. Additional categories for data collection should include: (1) teacher supply data at the state level; (2) the impact of teacher testing on recruitment, retention, and educational quality; (3) the nature of certification and licensure; (4) the experience of beginning teachers; (5) teacher development via analysis by years of service; (6) the prevalence and nature of career ladders; (7) teacher retention; and (8) early retirement.

AMERICAN LIBRARY ASSOCIATION
50 EAST HURON STREET · CHICAGO, ILLINOIS 60611 · (312) 944-6780

October 3, 1985

Emerson Elliott,
Administrator
National Center for Education Statistics
1200 19th Street, N.W.
Washington, DC. 20208-1401

Dear Emerson Elliott:

Please forgive ALA for not responding earlier to your letter of May 17, 1985 addressed to Robert Wedgeworth. Mr. Wedgeworth left ALA in August to become Dean of the College of Library Science at Columbia University. During the last few months a number of important things were neglected, your letter among them. We sincerely regret that ALA has not taken part in the redesign of the NCES data collection program for elementary and secondary education.

As you know ALA has been working actively with NCES to improve the collection and publication of statistics about libraries. Our Office for Research completed a contract for NCES in November, 1984 with a report entitled "Analysis of Library Data Collection and Development of Plans for the Future." This report recommended revisions in the forms used to collect statistics from College and University Libraries and from School Library Media Centers. We have been pleased to learn that both of those forms are being used this fall in much needed surveys. Robert Wedgeworth, ALA Executive Director, and Jo An Segal, Executive Director of ALA's Association of College and Research Libraries sent a letter to the directors of all college and university libraries urging them to complete the form promptly and completely. ALA's American Association of School Librarians is eager to see the results of the Fall 1985 Survey of School Library Media Centers as the data will be extremely useful to the committee engaged in drafting revised standards for school library media centers. Finally, the ALA Office for Research has just begun work on a contract, funded jointly by NCES and the Division of Library Programs, to conduct a pilot study leading to a cooperative system for public library data collection based on annual data collection by the fifty states.

As you can see ALA is very involved in the NCES data collection efforts which relate to libraries. School library statistics are of special concern to us because there is almost no other source of information on this topic. The OFR report noted earlier documents the fact that although some information about the other library types is available from states and other organizations, this is not true for school library statistics. We are very pleased that NCES is surveying school library media centers in Fall, 1985. The latest available data was collected in 1978 and is no longer useful. We believe that statistics about school libraries should be an important part of any elementary and secondary school data collection effort. Data should be collected regularly on this topic and included in the compilations which describe other data collected on education--the Digest of Education Statistics and/or the Condition of Education.

We hope these ideas can be incorporated into your plan even though we have missed your deadlines. If there is anything ALA can do at this late date to provide additional information please contact us through Dr. Mary Jo Lynch, Director of the ALA Office for Research.

Sincerely yours

Roger H Parent

Roger Parent,
Acting Executive Director

cc: Eileen Cooke, Director, ALA Washington Office
Larry Lamour, NCES
Mary Jo Lynch, Director, ALA Office for Research

RP:ld

EDUCATIONAL STATISTICS FOR STUDIES OF POLICY AND ADMINISTRATION

Robert G. Lehnen
Professor of Public Affairs
School of Public and Environmental Affairs
Indianapolis, IN 46223

Telephone: 317/264-3466

(August, 1985)

---

an invited paper prepared for the redesign of the elementary and secondary education data program of the United States Department of Education, National Center for Education Statistics

Education Statistics for Studies of Policy and Administration

I. Introduction

The observations contained in this paper arise from the study Financing Indiana's Public Schools done in 1984 for the Indiana General Assembly[1]. The study used data from the National Center for Education Statistics (NCES) to compare Indiana with other states in the areas of resources committed to education and in educational performance. The difficulties that arose in making these comparisons and the reactions of the supporters and critics of the study are reported here. The remainder of the paper addresses some specific problems concerning NCES statistics and practices and presents some recommendations for improving them.

II. Utility of NCES Statistics for Policy Studies

A central question facing most states today concerns the adequacy of their public education system, both in the areas of the resources committed to education and the performance of the system. Indiana is no exception in this regard. Having experienced severe economic hardships in many parts of the state, Indiana in recent years has taken a closer look at its public schools with the intent both to improve quality and also to make the state more competitive in its ability to attract industry and retain its workforce.

Financing Indiana's Public Schools was designed to review Indiana's position among the states and report on the effects of property tax reform, undertaken in 1973, on its 304 school districts. NCES statistics played a central role in accomplishing the first purpose[2], and statistical information .pn2 from the Indiana Department of Education provided the basis for district by district comparisons of the effects of tax reform. The discussion in this section is confined to the role that NCES statistical information played in the Indiana report.

The principal NCES measures used in the report fell into two categories: (a) measures of input (resources) and (b) measures of output (performance). The report reviewed the availability of education statistics that both measured, in some general way, one of these two concepts and also provided state by state comparisons. Most measures reported in NCES publications did not meet these two requirements, particularly the latter one.

The measures eventually used in the study to compare Indiana to other states are as follows:

Input Measures

Average days attended per pupil enrolled, 1975-76
Number of pupils per teacher, 1980-81
Number of pupils per teacher based on enrollment, 1980-81
Total expenditures per pupil
Total public school expenditures as a percent of personal income

Output Measures

Median years of education
Percent graduating from high school
Average SAT score (for 22 states)

Some of these measures were not ideally suited for the task. The output measures often were criticized as being too vague and not reflecting performance. In particular, the lack of suitable measures of achievement and other aspects of educational performance limited the effectiveness of the study.

Among the input measures Total Expenditures per Pupil received the most attention. A series was constructed from data provided in tables reported in issues of the Digest of Education Statistics to show trends in national and state expenditures. This analysis received considerable comment from General Assembly members, the media, and various interest groups. Those critical of the conclusions of the study--that Indiana was substantially behind other states in levels of spending for public education and was falling further behind--argued that the NCES data were unreliable and were not uniformly reported by the states. The critics thus concluded that Indiana was, in fact, better off than what the NCES statistics indicated, and the conclusions of the study must therefore be discounted. These and other issues in the use of NCES statistics for making policy recommendations are discussed in more detail in the next sections.

III. Improving the Utility of NCES Statistics in Policy Studies

Performance Measures: One conclusion reached in doing the report for the General Assembly is that there are few good measures suited for policy and administration studies. One can divide policy and administration measures into three categories: input or resource measures, process or administration measures, and outcome or performance measures. Although some measures reported in the Digest of Education Statistics may be suitable to use for one of these three purposes, most fail on other accounts discussed below.

The area where most attention is needed is on the performance side. The question most asked by Indiana General Assembly members was about the effects obtained from various programs and expenditures. How can one know that if spending is raised, or class size reduced, that the schools will produce a better educational product? The three output measures used in the Indiana policy study were poorly suited to evaluate the performance of the Indiana educational system as a whole, and thus anecdotal information received as much credibility at times as national statistics.

Less Aggregation: The current measures of education (input, process, or output) are not very useful because of their present level of aggregation. National averages and other statistics do not reveal much about the state education systems. Since education is primarily a local and state function, it is essential to disaggregate the information to state and district levels. Without this detail NCES data will have only limited utility for policy studies within states. Yet it is the states who will determine the direction and scope of education policy and not the federal government. Without this detail NCES data will have only limited utility for policy studies among or within states.

Comparability: It should be possible to obtain uniform information about every school district in the nation. Since most states including Indiana have their own departments for recording state and district level education statistics, one may argue that the state is the proper place to maintain such detail. If the states are to become the repository of state and district information, then the measures reported by the states and NCES must be the same. In the course of the Indiana school finance study, it was not possible to construct district-level measures of "national" statistics even though the information had come from Indiana. For example, the measure Total Expenditures per Pupil reported by NCES was not available by district in Indiana.

Lack of Documentation: There appears to be no technical publication reporting NCES operational definitions, technical terms, standards, practices, and quality control. Early in the Indiana school finance study, a reference librarian at Indiana University attempted to obtain such a document without success. Subsequent calls to NCES and a conversation with a staff member revealed that no such publication presently exists. The lack of such documentation makes it impossible to provide information about the interpretation of the statistical information. Furthermore, it compromises the conclusions reached by analysts using NCES data, because critics often use anecdotal or hearsay information to refute conclusions. For example, critics of the Indiana study charged that Indiana's average expenditure figure was "too low" because book fees, paid for in Indiana by parents and not by tax monies, were not included in the Indiana statistic. This observation could never be verified or refuted.

Media for Reporting: The eight NCES measures used in the Indiana study came from the following sources: key punching of selected tabular information from various issues of the Digest of Education Statistics, and key punching of a table reported in USA Today from the January 6, 1984, news conference by Secretary of Education Terrell Bell. The latter source was subsequently verified six months later by obtaining a copy of the press release prepared by the Department.

Several comments are in order regarding the means by which this statistical information was obtained. First, statistical series, such as total expenditures per student, should be identified as such and reported as a series. Presently, finding comparable tables in various issues of the Digest appears to be the only direct method of identifying series. Second, data should be available through other media than press reports and publications. The data released to the media during the Bell press conference should have been available on floppy diskette, computer tape, and other such media. In general, Digest issues should appear on computer tape in a manner similar to that used by the Bureau of the Census to report its County-City Data Book information. With the advent of the professional computer, data should be readily available on floppy diskette as well as on public-use tapes.

As a case in point, the Indiana Department of Education provided its entire 10-year database for 304 school districts on computer tape in Statistical Analysis System (SAS) file format. Such assistance greatly facilitated the access to information and the analysis of the issues.

In the next section suggestions are presented for the improvement and/or expansion of educational statistics. This list is not designed to be comprehensive but rather reflects the accumulated experiences and opinions of the author.

IV. Some Specific Suggestions Concerning Education Statistics

Bureau of the Census Data: The Bureau of the Census presently provides data on population and housing characteristics by school district. This series is a important source of information and should be continued. Two observations pertain to the Indiana public-use tape. First, the release of data should occur sooner. Secondly, the accuracy of the data needs to be verified. In the latter situation, the Indiana public-use tape contained numerous errors, including omission of districts and the combination of similar-named districts, and thus was unusable for the study.

Public Opinion toward Education: Several polling organizations such as the Gallup Poll have conducted surveys of public attitudes toward education issues. Such surveys should be continued and coordinated through NCES. Specifically, a standard national survey of opinion should be conducted through NCES at least annually and a series of standard indicators developed. In addition, NCES should have a research program whereby specific questions may be added to the core survey to measure current issues. For example, the impact of private schools such as the "Christian academies" on public schools could be explored. The survey data should be available in a timely manner on public-use tape or similar medium. Competitive research solicitations should be offered to select the principal investigator for each year's special subject.

Survey of School Personnel: No doubt one of the least documented areas of education is the state of the education personnel system, the teacher and administrative workforce. One reads about "burn-out", victimization, and disillusionment among teachers and administrators, but NCES provides little in the way to document these reports. An annual survey of school personnel, including at least teachers and administrators, is needed to measure the conditions in the workforce. Some topics to be included in the survey are measures of "burnout", perceptions of working conditions, reports on use of summer time, expectations about the future, crime and victimization in the workplace, out-of-pocket expenses incurred by faculty, and uncompensated job-related duties.

Labor Relations Information: NCES needs to report on the state of labor relations in the nation's schools. It should provide information on such labor-related characteristics as the degree and type of professional bargaining units, the number and duration of strikes, and the time to settlement of contract negotions. These data should be available on a district by district basis. One question that might be ultimately answered from such data is the relationship between labor relations (timely, amicable settlement vs. strife-ridden negotions) and performance.

Health and Nutrition: Many people believe that the nutritional behavior of students, both before school and in the school cafeteria, is related to discipline and other performance issues in the classroom. School cafeterias vary greatly in the degree to which they offer nutritionally balanced meals as contrasted to ones high in carbohydrates and "empty calories". Some teachers have observed what they believe to be "carbohydrate highs" that may cause behavior problems in afternoon classes.

NCES, possibly in cooperation with the Department of Health and Human Services, should develop a series of studies to determine the nutritional value of meals offered in school cafeterias and consumed by most students. In addition, the nutritional education and behavior of students should be explored, including breakfast-eating behavior, knowledge of nutritional issues, and choices make in the cafeteria, and this behavior should be related to educational performance and behavior.

Class Size and Teacher Load Information: The current measures of average class size reported in the Digest of Education Statistics do not provide sufficient detail to be of much use. The averages reported for Indiana, for example, in no way reflect the personal experiences of this author or those of teachers he has consulted. One general argument made locally is that special education classes skew the class size distribution and distort the mean, thus giving the impression of smaller than actual class sizes. NCES needs to develop information on the variation in class size by distict, subject, and grade. An example of class size data is given in Table 1. Other statistical information such as the median class size for each subject and percentile information should be developed from such data.

TABLE 1

Example of Class Size Table

School District: ABC School Corporation
Grade: 7

Size of Class:	18	19	20	21	...	Median
Subject						
English						
Mathematics						
Reading						
Science						
Social Studies						

table entries are number of classes of a given size for each subject

TABLE 2

Example of Teacher Class Load Measure

School District: ABC School Corporation
Level: Middle School

Number of Students:	120-129	130-139	140-49	...	Median
Subject					
English					
Mathematics					
Reading					
Science					
Social Studies					

teachers who teach part-time or more than one subject should be pro-rated on an FTE basis

A related measure that should be developed is Teacher Class Load. This is a measure of how many students a teacher in a given subject instructs each school day. Such data should be available by district and subject. An example of such data is given in Table 2. The development of the above information should be done for core subjects initially and then expanded to include the entire cirriculum.

Measures of Performance: As indicated in the first section of this paper, NCES does not report sufficient measures of performance. The general question of what kinds of measures to be collected should be explored with various public interest groups, policy makers, and education professionals. Undoubtedly, some measures of achievement in basic subjects are required, but the measures should not be confined to achievement measures. The selection of measures to be included in NCES reporting should be done by a public process reflecting the contributions of many diverse groups. Whatever measures are selected should be reported at least annually by state and district. Without such information the nation's policy makers cannot effectively evaluate the nation's schools and develop programs to remedy deficiencies.

Comparable Expenditure and Revenue Data: The lack of comparability between states poses serious problems for understanding that nature of school expenditures and revenues. Although sufficient detail exists within Indiana for its 304 school districts, attempting to compare Indiana's practices to other states is extremely difficult or impossible. NCES should take the lead in developing a model state data base and reporting system for district level data. Although such data may be collected and maintained at the state level, standard format public-use tapes from each state should be available.

Expenditure and revenue measures should be the core indicators of such a system but other measures such as enrollment and performance measures should be considered as well. The separate states may take responsibility for collecting and reporting the information, whereas NCES may report statistics of primarily national interest.

V. Some Concluding Observations

The present condition of NCES statistics severely limits their utility for policy and administration studies. Although this paper has suggested several areas where improvement is desirable, it should be noted that some recommendations have special priority. The principal areas for improvement should concentrate on developing more useful measures of education performance; producing less aggregation of information by providing state and district level information; and finally instituting better documentation, quality control, and distribution of the product. These enhancements, more than any other, should improve the condition of national education statistics.

Footnotes

[1] Robert G. Lehnen and Carlyn E. Johnson, Financing Indiana's Public Schools: An Analysis of the Past and Recommendations for the Future. Indianapolis, IN: Indiana University, School of Public and Environmental Affairs (1984), 160 pages.

[2] See "Chapter 4: Achieving Quality Education in Indiana: What Level of Funding Is Required?" in Lehnen and Johnson, op. cit.

Houghton Mifflin Company

One Beacon Street, Boston, Massachusetts 02108
(617) 725-5000 Cable HOUGHTON

School Division

June 20, 1985

Mr. Emerson J. Elliott
Administrator
National Center for Education Statistics
400 Maryland Ave. SW
Washington, D.C. 20202

Dear Mr. Elliott:

Thank you for your letter to Mr. Townsend Hoopes, President of the Association of American Publishers, inviting our industry to make suggestions for the planned NCES redesign of its elementary and secondary data program. The AAP views this as an exciting opportunity and has asked our Research Committee to respond.

At the June meeting of our committee, we agreed to do this in stages: (a) by filing with you, by June 21, 1985, a written list of suggested changes in NCES data collections for your published reports; (b) informal discussion of these suggestions with staff members of NCES during our committee's planned visit to your offices on August 1, 1985; and (c) if it appears warranted, to follow these steps with further written communication by September 30, 1985, and/or participation in public hearings, as mentioned in your letter to Mr. Hoopes. The following is our committee's list of suggestions:

1. Estimates of secondary school course enrollments are needed much more frequently. By this we mean the type data NCES has supplied Summary of Offerings and Enrollments in Public Secondary Schools, 1972-73 (NCES 76-150), Course Offerings, Enrollments, and Curriculum Practices in Public Secondary Schools, 1972-73 (NCES 77-153), and A Trend Study of High School Offerings and Enrollments: 1972-73 and 1981-82 (NCES 84-224). This is a critical data need. Such market-size estimates comprise one of the most vital factors by which educational publishers decide whether to publish and how to publish instructional materials. Such data every 10 years is clearly not frequent enough. Dramatic changes occur in a decade. Elhi publishers need such data every 2 years. We submit that sufficiently reliable data can be collected through probability samples at a reasonable expenditure by the Government. Enrollment data should cover Grades 7-8 as well as Grades 9-12. Reasonably reliable sub-sample estimates should show enrollment variations for (a) course duration (full-year, one semester, etc.), (b) geographic distribution (state by state, or perhaps by the nine census regions), and (c) public vs. private schools. The raw numbers of students enrolled in each course, plus their percent of all students in all grades, are the key data needed.

2. Estimates every two years of elementary school enrollments for each of the major course areas are likewise a critical need. This should cover Grades K-6, especially if you account for Grades 7-8 enrollment components of K-8 curricula under Item #1 above. Experts presume that 100% of students are enrolled in Reading and Mathematics in Grades K-6. However, as we know, there is considerable variation, especially in the lower grades (as well as in Grades 7-8), in the proportions of students enrolled, at each grade level, in other course areas such as: English/ Language Arts, Spelling, Handwriting, Science, Health, Social Studies, Computer Sciences, Music, Art, Foreign Language, etc. Yet there is currently no central data collection pinpointing the sizeable variations believed to exist in the percents of students, grade by grade, who take these subjects. This is a serious data gap. The same data specifications (especially concerning the duration and/or frequency with which such courses are taught) -- and the same supporting arguments -- outlined in Item #1 above apply here.

3. Grade by grade projections of total enrollments (of all students in all courses) for each level K-12 are a similar vital and frequent need for long-range planning by publishers. Data should be organized like Table 6 in the NCES volume, Projections of Education Statistics to 1990-91, an important annual document. Past data (from Table 22 in the NCES annual Digest of Education Statistics, another important document) should extend back 10 years, and projected estimates should extend 10 years into the future. Such projections could be delivered in both of two frequencies:

   (a) Annually: Grade by grade, K-12, for nationwide total enrollments, as well as sub-sample breaks for (1) the nine census regions and (2) public vs. private schools.

   (b) Every 2-3 years: Grade by grade, K-12, for nationwide total enrollments, as well as sub-sample breaks for state by state.

4. The number of units required for high school graduation in each of the various course areas, state by state, are important data -- needed annually.

5. The specific courses (and their duration) mandated in high schools, state by state, are similarly important data -- needed annually.

Our Research Committee, Mr. Elliott, respectfully submits that the above key data needs will enhance decision-making on a broad basis throughout the education community. More prudent decisions by publishers, large and small, as represented by the AAP, lead to a better choice of more competitive and suitable instructional materials for all school systems.

Please let either me or Ms. Barbara L. Meyers, Assistant Director, School Division, AAP, at its headquarters in New York know whether it will be convenient for our Research Committee to discuss these and perhaps other issues with you and NCES staff members at our scheduled meeting in your offices August 1 (as arranged by her and Ms. Kay McKinney of NIE).

Sincerely,

Berry L. Reece
Chairman, Research Committee
School Division, AAP

cc: Ms. Barbara L. Meyers, AAP
Mr. Donald Ecklund, AAP

THE ASSOCIATION OF TEACHER EDUCATORS

THE IDENTIFICATION OF DATA NEEDED TO SUPPORT DELIBERATIONS ON POLICY ISSUES

Submitted to the National Center for Education
Statistics for consideration in the redesign
of its elementary and secondary education
data program.

Prepared by Dr. Lee Bartolini

The Association of Teacher Educators, as an organization concerned about the improvement of elementary and secondary education, is keenly interested in the National Center for Education Statistics' plan to redesign a 10 year program for elementary and secondary data collection. The organization is primarily interested, however, in data collection activities which will focus on the needs of teachers and teacher educators. In recent years, a flood of national and state reports have identified the need to improve the quality of teaching as a major educational issue. Data collection activities designed to provide the information regarding this issue would be most beneficial to decision and policy-makers.

Some institutions, primarily state education agencies, have tried to identify and collect data needed to improve the quality of teaching and teachers. These efforts have included, but have not been restricted to, the components of teacher training programs. Additional information useful to decision-makers includes local district recruitment, selection, and evaluation procedures, and statewide teacher supply and demand statistics. These subjects are areas in which new data will be especially useful to decision-makers for elementary and secondary programs.

The Council of Chief State School Officers (CCSSO) is also interested in knowing what states are doing to improve the quality of teaching. A committee of CCSSO on Teacher Education and School/College Collaboration recently initiated a 50-state survey which attempts to identify those activities which have occurred or will occur relevant to four policy areas: attracting persons to the teaching profession, preparing persons for teaching, licensing persons for the teaching profession, and retaining teachers. The CCSSO survey also includes a component on teacher supply and demand. The policy areas identified by the CCSSO and the state education agencies provide a framework for identifying data needs.

The NCES plan to redesign its data collection activities focuses upon elementary and secondary education. However, as suggested above, some data identified may be available from or through state education agencies or regional levels of government. NCES may want to consider these sources when planning data collection activities. Another consideration is that teacher education, historically, has been a state concern. Therefore, national data would be most useful if it could be generalized to specific states.

Much information needed to make decisions regarding teaching is available only through colleges and universities. Information on the recruitment and selection of potential teacher candidates by institutions of higher education, information regarding the components of teacher education programs, and information concerning requirements for satisfactory completion of teacher preparation programs are examples of data needs which greatly affect elementary and secondary education, but which must be obtained through institutions of higher education. NCES also needs to consider this source in its data collection plan.

Specific types of data relevant to the improvement of teaching in local schools are outlined in the following sections. Consistent with the scope of the NCES plan to redesign its data collection activities, the data needs identified focus primarily, but not exclusively, on elementary and secondary education. Data have been identified by posing a series of questions. These questions have previously been raised by policy-makers and serve as guidelines for identifying specific data which need to be obtained through data collection activities.

A SAMPLE OF
QUESTIONS TO BE USED TO IDENTIFY
DATA NEEDED TO SUPPORT DELIBERATIONS ON POLICY ISSUES

I. Teacher Supply and Demand. Teacher demand is usually defined in terms of needed staff in specific subject areas within individual states. Colleges and universities, however, prepare teachers not only for the states where they are located, but for the rest of the nation as well.

- A. What is the supply and demand balance for teachers by state, by region, and by specific subject area?
- B. In what states or regions is demand expected to increase? Decrease?
- C. How difficult is it for teachers prepared in one state to move to another state where need might be greater?
    - 1. What certification requirements are common to most states?
    - 2. What core of certification standards and qualifications would enable persons preparing to teach to meet most state requirements?
    - 3. How common are reciprocal agreements between states which would allow persons prepared in one state to teach in another?
- D. Can non-teacher experts (persons not prepared as teachers but recognized as experts in subject matter areas) be used as classroom teachers? What are the restrictions or limitations?

II. Recruitment of Prospective Candidates into the Teaching Profession. Much has been said about the increased opportunities for women in the field of business and other professions, thus removing them as potential teacher candidates. There is also concern that the teaching profession does not attract the best and brightest students.

- A. Why do persons choose or not choose to become teachers?
- B. What local working conditions, salaries, or social conditions would make teaching more attractive to prospective candidates?
- C. What problems are associated with recruiting staff for extracurricular activities?

III. Recruitment and Selection of Teachers by Local Districts. Local districts may use a number of techniques to recruit and/or select teachers. While some criteria are well established, such as a satisfactory academic record, little is known about the variety of criteria or degree of difference in the use of specific criteria. In addition, little is known of the factors which limit or constrain recruitment and selection practices.

- A. What procedures are used by districts to recruit teachers? What techniques have been particularly useful in identifying and recruiting quality candidates?

B. What do local districts do when qualified candidates cannot be recruited?

C. What practices are employed to recruit minority candidates, in addition to routine recruitment practices?

D. What factors most constrain recruitment of qualified personnel in local districts?

E. Are district selection procedures clear and well defined?

   1. Do districts have fully developed job descriptions?

   2. What qualifications are required of all candidates?

   3. Do district qualifications for positions exceed minimum state requirements?

   4. Does the district (or state) require candidates to take qualifying tests? What tests?

   5. Who are the personnel who actively participate in the selection of educational personnel?

   6. What are the most important criteria used in the selection of teacher candidates?

   7. What does a district look for during an interview with a candidate?

   8. Does the district assess a candidate's writing skills or abilities?

   9. Do districts routinely select applicants who are certified to teach in more than one area?

IV. Evaluation and Performance of Teachers. Evaluation practices implemented in local districts may have an important effect on the quality of education. More needs to be known about how staff evaluations are conducted and what practices, if any, are employed in staff development. There is also concern that budget restrictions and declining enrollments have caused school administrators to assign tenured staff to teach in fields for which they lack sufficient preparation.

A. Do districts employ formal evaluation procedures when assessing the performance of teachers?

   1. Are standardized evaluation instruments used?

   2. What specific criteria are used to evaluate staff?

   3. How often are staff evaluated?

B. Who, in local districts, evaluate teachers? What special training is required of these personnel?

C. If an evaluation indicates that a staff member's performance is unsatisfactory, what subsequent action is taken?

D. What programs for staff development currently exist in local schools?

E. What restrictions or constraints hinder staff development or the improvement of performance?

F. What relationship exists between performance on standardized teaching tests and performance in the classroom?

G. What evidence is there to suggest that persons, either because of insufficient formal training or because of a long absence from a given teaching field, are being asked to teach classes for which they are inappropriately prepared?

V. Retaining Teachers. Major concerns of those interested in improving the quality of education are retaining the best and brightest practicing teachers and removing the incompetent teacher. Conventional wisdom suggests that many of the best teachers leave the profession for positions in private industry. Yet, little is known about those who leave.

A. What is the attrition rate of teachers? Is teacher turnover greater or less than turnover of personnel in other professions?

B. Are those who leave the teaching profession the most qualified teachers? The better performing teachers?

C. Why do practicing teachers leave the profession? If they leave, do they ever return?

D. What working conditions or approaches (e.g. merit pay, differentiated staffing, etc.) would act as incentives for keeping the best teachers in the classroom?

   1. What are the factors that provide the most job satisfaction for teachers?

   2. What are the factors that create the most dissatisfaction for teachers?

E. What are the constraints associated with retaining the most competent teachers?

F. What are the constraints associated with removing the incompetent teacher?

COMMENTS

OF THE

COUNCIL OF CHIEF STATE SCHOOL OFFICERS

ON THE

NATIONAL CENTER FOR EDUCATION STATISTICS

REDESIGN

OF THE

ELEMENTARY AND SECONDARY EDUCATION DATA PROGRAM

The Council of Chief State School Officers is pleased to have the opportunity to comment on the National Center for Education Statistics ten year plan for elementary and secondary data collection. The Council is a non-profit organization comprised of the state superintendents and commissioners of education in the fifty states, six extra-state jurisdictions and the District of Columbia. They are the executives responsible for administering the nation's public education enterprise consisting of approximately 85,000 schools with approximately 40,000,000 students and an annual expenditure exceeding $119,000,000,000.00. When combined with postsecondary and continuing education, this enterprise requires 37.8% of state government expenditures, 42.1% of local government expenditures, and combined with federal contributions (4.1% of federal expenditures) total expenditures for education rank second only to National Defense and International Relations in terms of expenditures from all levels of government.

It is natural, then, that the stewardship of this enterprise demands complete and accurate information for accountability to the public and legislative bodies, for the support of effective decision-making, and for the assessment of educational progress. As a result, the Chief State School Officers as collectors, processors, responders, and users of education statistics are in a unique position to provide insight into the implications of plans, changes and needs for data and information about education.

The Council believes that the National Center for Education Statistics has a vital role in responding to educational needs in the following general areas:

1. Contextual Parameters or descriptors that describe the educational enterprise
2. Indicators of the health or status of education and its relationship to other countries
3. Special longitudinal and other statistical surveys and systems that are practical only at the national or federal level
4. Assistance to state and local agencies in the design and operation of activities at the state and local level.

An effective combination of these four areas will result in increased opportunity for new information to be generated by the Center, researchers using Center data, and by policymakers analyzing effect and impact of change.

It is important to point out at the onset that the degree to which these four responsibilities can be appropriately met is highly dependent upon the level of funding for the various activities. The Council at is November 1984 meeting stated that the U.S. Department of Education should "Request increased appropriations for assessment and evaluation efforts by five to six times the current level (8+ million per year) to make the capability comparable with national reporting in health, agriculture and other federal statistical functions. "A failure to accept the cost of producing, reporting, and analyzing statistical information, and the subsequent provision of funds to support this cost, will limit any real advance to piece-meal efforts with neglible improvements."

1. Contextual Parameters or Descriptions

Although a great deal of attention has recently been focused on the combination of data elements to assist in policy analysis, there will always be a need for basic education data. How many districts, schools, students, teachers, administrators, etc. are there? How much money is being spent? Data that define the size and the scope of the enterprise are essential and will continue to be needed by users.

The Council and its Committee on Evaluation and Information Systems (CEIS) has had a long and consistent history of supporting the concept and implementation of a Common Core of Data that describes the system statistically. The philosophy of a federal-state cooperative data system to respond to federal data needs from state and local administrative records is one that has been pursued since 1961. The necessity of maintaining a cooperative approach is critical to the continued and enhanced ability of the federal government to collect either voluntary or contracted data on a systematic basis. Consequently, we applaud the National Center for its approach in soliciting input from a broad variety of audiences on a formal basis. We encourage the involvement of CCSSO and its Committee on Evaluation and Informatlion Systems at each stage of this process.

CCSSO encourages the Center to cooperatively define those data elements that can be efficiently collected with universe information that will improve sample selection procedures without unduly increasing reporting burdens. This would include data on the school district universe and the school universe.

Standardization and coordination of data definitions at the federal level is a role that may be appropriate for NCES. This coordination, and the attendant acceptance of the development and distribution of glossaries, is necessary to promote the improved comparability of information. Additionally, the acceptance of this role would increase the confidence of data users that information in given formats would be available over time and not subject to changing program emphasis or approaches.

The Council will gladly assist, through its CEIS as well as in other appropriate ways, in the identification of useful, necessary, or improved CCD data elements as well as suggesting elimination of those that have proven to be of little value.

The Council looks to NCES as a provider of information relative to non-public schools as a basis for analyzing total educational information. Additionally, the establishment of comparable statistics about education in other countries would be most useful as states analyze their own data sets.

Finally, the provision of current information is a goal that NCES should constantly be striving to improve. The CCSSO recognizes the problems in collecting data and its impact on the delay in publishing information but is convinced that a rapid turnaround of information is essential for improved services. The recently established bulletin board is a positive step in this direction but electronic display of old information is holding out only half a promise for improvement.

2. Indicators.

The Council fully endorses the *Condition of Education* and the *Indicators of Education Status and Trends* and encourages continued cooperative development in conversion of statistical data into information that is useful in describing the effects of the schooling process. Appropriate statistics that are not part of the Common Core of Data should be gathered by NCES from other sources or through special surveys or procedures using sampling whenever possible. Analysis of the design of these special surveys or activities should consider the possibility of state use and in addition to the national requirements. The aggregation of data about education collected by other federal agencies (such as those reported in the *Condition of Education* and the *Indicators*) into a common, accessible data base such as the newly created bulletin board, could be of considerable value to the states as data users.

3. Special longitudinal and other statistical surveys and systems that are practical only at the national and federal level.

The Council is fully supportive of the NCES High School and Beyond Survey and the planned National Education Longitudinal Survey of 1988.

The activities have proven to be extremely helpful in generating a variety of new information on courses, attitudes, relationships, and results. As stated earlier we would encourage consideration of developing these activities in a way that results in state representative as well as nationally representative data. It is recognized that this introduces additional cost for these activities but CCSSO looks to NCES to serve as the state's advocate in soliciting necessary funding to accomplish this.

4. Assistance to State and Local Agencies

A review of effective statistical and information systems clearly indicates that successful programs are dependent upon the capability of respondents to provide accurate information. Accurate and reliable information at the federal level is possible in direct proportion and relationship to the development and improvement of support systems at the state and local level. This concept is embodied in the federal-state cooperative data collection systems which have involved direct federal financial assistance to states for the development of their systems. Such cooperative systems have become operational in the Department of Labor, Bureau of the Census, Law Enforcement Assistance Administration, and the Department of Agriculture. Other specific examples of such systems incude the Coopeative Health Statistics Systems and the Medicaid Management Information System.

In education, however, assistance activities have had an uneven history and have been a woefully underfunded. As resources from all levels become more scarce the competition for these limited funds by all aspects of general program administration will make it more difficult to develop improved data systems. Information and statistics are not a natural by-product produced at no cost; but rather a commodity which must be produced by someone and paid for by some agency. The CCSSO encourages NCES to seek sufficient resources to permit all levels of government to have resources to generate the data that the Federal Government needs to report timely, accurate and comprehensive statistics. If the resources are not made available, inadequate reporting results or funds and efforts must be diverted from more important activities related to program administration or instruction. Neither of those alternatives is acceptable.

As indicated earlier, the Council would be most happy to assist in the detailed development of specific data items that make up the component parts of the ten-year plan in a more thorough manner. The need to analyze use of data collected, as well as the cost of data to provide are questions that need a great deal of attention and assistance from state and local providers and are as important as the definition of data elements. A federal data system that is not useful or practical to the state and local education agencies providing information will not succeed. Effective dialogue, coordination and assistance will allow an enhanced opportunity for all partners to access and use valid, reliable and timely statistics.

President
GORDON M. AMBACH
New York Commissioner
of Education

President-Elect
FRANKLIN B. WALTER
Ohio Superintendent
of Public Instruction

Vice President
FRANK B. BROUILLET
Washington Superintendent
of Education

Directors
VERNE A. DUNCAN
Oregon Superintendent
of Public Instruction

JERRY L. EVANS
Idaho Superintendent
of Public Instruction

DAVID W. HORNBECK
Maryland State Superintendent
of Schools

STEPHEN KAAGAN
Vermont Commissioner
of Education

WAYNE TEAGUE
Alabama Superintendent
of Education

CAROLYN WARNER
Arizona Superintendent
of Public Instruction

Executive Director
WILLIAM F. PIERCE

July 19, 1985

Mr. Leslie Silverman
Deputy Assistant Administrator
Division of Statistical Services
National Center for Education
Statistics
U.S. Department of Education
1200 Nineteenth Street, N.W.
Washington, D.C. 20208-1401

Dear Leslie:

At our July 16, 1985 meeting of the CCSSO Ad Hoc Committee on the NCES Elementary-Secondary Data Program Redesign Project, it became apparent that the direction and limits of the project would be impacted by the perceived mission and functional boundaries assumed for the National Center for Education Statistics. We strongly urge that the function be a true statistical center that assumes the major responsibility for coordination of the collection, assembly, analysis and dissemination for that sector of society under its purview, namely education.

The Secretary of Education would be required to make a clear and committed designation that the Center would have responsibility for coordination of statistical data collection and analysis activities across the Department of Education regardless of organizational lines and/or bureaucracies. This assignment would also require that the Center be charged with promoting the integration of the numerous data collection activities conducted by other federal agencies (Department of Agriculture, Bureau of the Census, Department of Labor, et al.) and related private agencies (National Education Association, American Council on Education, and the testing industry) to minimize burden on respondents and to develop increased standardization of terminology.

Education ...a sound investment in AMERICA.

COUNCIL OF CHIEF STATE SCHOOL OFFICERS
379 Hall of the States, 400 North Capitol Street, N.W., Washington, D.C. 20001 • 202/393-8161

This coordination role would include: 1) first and foremost, the coordination of the various activities currently under development in NCES (e.g., CED, VEDS, NELS-88); 2) expansion of the system to include those other data collection activities by the Department of Education (e.g., Special Education, Chapter I of ECIA, Chapter II of the Math and Science Act); and finally 3) establishment of out-reach activities to other agencies to ensure appropriate federal and national coordination. Included in this function would be defining a common set of data elements across the spectrum, coordinating collection of all statistical data, developing efficient collection and dissemination systems (in conjunction with users and providers), seeking out current needs for educational information, and providing assistance, both technical and financial, to the respondees and users of educational data.

Any effort at a ten-year plan, without a clear understanding of the agency's mission and philosophy, offers little promise of success. Additionally, in our view, the failure to expand the mission and functional boundaries of the National Center to a true center for education statistics limits the potential growth to little more than that capacity which exists today.

Sincerely,

George Rush

George Rush
Staff, Council of Chief State
School Officers

GR:fkc

cc: Emerson Elliott

CHIEF STATE SCHOOL OFFICERS • COUNCIL •

President
GORDON M. AMBACH
New York State Commissioner
of Education

President-Elect
FRANKLIN B. WALTER
Ohio Superintendent
of Public Instruction

Vice President
FRANK B. BROUILLET
Washington Superintendent
of Education

Directors
VERNE A. DUNCAN
Oregon Superintendent
of Public Instruction

JERRY L. EVANS
Idaho Superintendent
of Public Instruction

DAVID W. HORNBECK
Maryland State Superintendent
of Schools

STEPHEN KAAGAN
Vermont Commissioner
of Education

WAYNE TEAGUE
Alabama Superintendent
of Education

CAROLYN WARNER
Arizona Superintendent
of Public Instruction

Executive Director
WILLIAM F. PIERCE

September 20, 1985

Emerson Elliott
Administrator
National Center for Education Statistics
1200 19th Street, N.W. #606D
Washington, D.C. 20208

Dear Emerson:

The Council of Chief State School Officers's Ad Hoc panel on the NCES Elementary/Secondary Education Redesign Project appreciates the opportunity you have afforded the task force to provide input into this important review and planning process. The recent heightened interest in educational statistics and information for program reform, system accountability, policymaking and applied research suggests the need for an accurate, timely and comprehensive data base of statistics aggregated in a manner that does not place an undue burden on current local and state information systems. We are encouraged that NCES has initiated such a thorough review and look forward to assisting the Center formulate alternative approaches to address this need.

The Ad Hoc panel met on September 10th to review a draft of the "Synthesis of Invited Papers" and to consider future steps in the redesign project. This meeting resulted in the following general recommendations.

1. The Center should develop a clear mission statement, along with an organizing theory for the integration of administrative record systems, sample surveys and longitudinal activities. The statement should address the Center's role in approving, coordinating, aggregating, maintaining and reporting information collected about education from other units of the Department of Education and other Federal agencies.

2. The regional public hearings, which the Ad Hoc Committee supports as an effective means for fully involving many participants in the process of providing and using educational data, should be scheduled for the first quarter of 1986 rather than the last quarter of 1985. The delay would be justified by the time required to fully involve all parties impacted by these

Education
...a sound
investment in
AMERICA.

COUNCIL OF CHIEF STATE SCHOOL OFFICERS
379 Hall of the States, 400 North Capitol Street, N.W., Washington, D.C. 20001 • 202/393-8161

proceedings and supply them with appropriate information in advance, including drafts of the plan.

3. NCES should consider state representative samples on all NCES-sponsored surveys and longitudinal study activities. Although this approach could result in increased costs to the federal government, recognition of the potential importance of resulting information would justify the expenditure.

4. Any attempt to construct model state and local information systems should include an examination and analysis of the record of past efforts such as the Midwestern State Educational Information Project, the USOE Handbook Series, the Belmont Project, the Committee on Educational Data Systems Manual and others.

5. NCES should exercise caution in balancing the legitimate desire of researchers and policymakers for detailed information with the cost and capability of institutions providing information. In its redesign project, NCES should consider factors such as the separation between research and statistics, state and local policymaking as contrasted with the federal role, and finally, the cost of information systems and their potential intrusion on the instructional process.

6. NCES should approach collection of data directly from local agencies with caution. While this is appropriate at times for sample surveys, it intensifies and compounds extant problems of data definitions, comparability, reliability and potentially detracts from the possibility of developing administrative record systems that will meet a variety of needs. Additionally, appropriate federal/state/local protocol should be honored in intergovernmental communications. CCSSO has long recognized the need to collectivly work with the federal government to ensure that data collected is valid, useful and collected with a minimum of intrusion. The Committee on Evaluation and Information Systems (CEIS) continues to be an effective vehicle for accomplishing this task.

Again, the Ad Hoc Committee appreciates the opportunity of providing input into the process and encourages the writing team for the plan, NCES staff, and yourself to call upon us for assistance.

Sincerely,

George Rush

George Rush

ASSESSING THE EDUCATION STATISTICS INFORMATION NEEDS OF NON-SEA PUBLIC POLICY DECISION MAKERS

Invited paper prepared by The Council of State Governments'
Office of Information Services for the
National Center for Education Statistics.

by

Mr. E. Norman Sims, Director
Office of Information Services

Dr. Deborah A. Gona, Coordinator
Survey Research Services

June 1985

Assessing the Education Statistics Information Needs of Non-State Education Agency Public Policy Decision Makers

The last half century in the United States brought a significant change to education policymaking as our system of public education moved from one with a "political" character to one with a "professional" character. This has been noted by Cremin who explains:

> The schools of a century ago -- party-dominated, patronage-controlled, professionalism ignored -- were transformed over the next half century by a combination of businessmen, professionals, and a new breed of university-trained administrators. Their enemy was the political machine and the political boss.[1]

One outgrowth of this movement was the development of a distaste by the professional education community for all things political. As a result, the school profession began to maintain the "purity of its motives and values and the sinfulness of party activity and partisanship."[2] In the late 1950s, however, educators began to see themselves as the focus of a discontent generated by their clientele. In the 1970s, this discontent had grown to the point that serious questions were being asked about the quality and quantity of professional educational services which seemed to be growing in expense.[3] Most recently this concern has been expressed by the report of the National Commission on Educational Excellence which recommends a wide range of educational reforms to halt the "rising tide of mediocrity."

In part because of this discontent, there is now a movement at the policymaking level away from technical, professional educators toward domination by political actors. Social and economic conditions, which have in the past supported the image of professionalism and independence, have changed. Educational policymaking has become more political than technical.[4]

Accompanying this movement have been two other trends which are of equal or superior importance: a shift of the primary political arena for education from the local level (the local education agency, or LEA) to the state level (usually the State Education Agency, or SEA)[5]; and a resurgence of the non-SEA state decision makers, such as the state legislature and the governor, as major factors in the development of state educational policies.[6]

Various forces are combining to cause these shifts. They include: judicial actions -- particularly in the area of school finance -- which have forced many state governments to reconsider fundamental educational policies formulated within SEAs and LEAs; federal involvement in educational finance and policymaking which has also spurred the development and expansion of the SEAs; and the impact of movements to reduce taxes, particularly by those who felt that they carried an unnecessary local property tax burden.[7]

As educational policymaking has moved from the local, technical-issue level to the state, public policy decision maker, an unfortunate schism developed at both the level of the user (between the professional educator and the political decision maker) and the provider (between information and policy centers serving legislators and executive branch agencies and those serving educators).

Unfortunately, for those involved in improving state educational decision making by providing adequate information, this schism has more than historical importance and is nowhere more apparent than in the area of research into the information needs and uses of these state decision makers and the provision of information to them. Obviously the non-SEA political actors make important education decisions. But we know very little about the information they use, want or need in the process.

The federal government, through the U.S. Department of Education and its National Institute of Education and National Center for Education Statistics, has made great strides in helping to improve decision making and use of information in SEAs and LEAs, not only through research into better education programs but also through the the dissemination of the results of these research efforts. Indeed, numerous studies have supplied information about educational decision making within the SEA.[8] Moreover, work by a variety of educational groups has enlighted us about several aspects of information delivery to decision makers within the SEAs and LEAs.

The state Capacity Building (CBG) and Research and Development Utilization (RDU) projects, as well as the Research and Development Exchange (RDx) and National Diffusion network (NDN) programs, have provided crucial data about facets of the information process.

Throughout the working life of each of these projects, new knowledge emerged about the process of decision making, resource delivery and client assessment.[9] But as Mattas and Rawnsley have suggested, in the design and operation of information services it is important to know more than which members of the educational community (the direct clientele of the educational research community) make use of services offered and what information they request. Research interest should also be directed toward all knowledgeable, and unknowledgeable, information users and non-users who have influence on the policymaking process.[10]

Again what is quite clear as we consider the great national debates over such issues as the role of the private sector in education and the operation of our intergovernmental system, is that the major policy decisions which affect education will not be made entirely in the SEAs and LEAs. They will be made by legislators, legislative staffers, governors, budget directors, state planning officers and others. These other actors may be familiar with the results of educational research and its application to policy, but it is more likely, as we shall see later, that their knowledge is scanty (coming to them secondhand) and based upon information which is poorly provided or limited in applicability.

It is also likely that educational decisions at the state level will be based upon information provided by sources other than those normally considered as information providers by the education community. The Council of State Governments 1984-85 edition of the Book of the States, for example, lists 90 organizations which may be called upon by state policymakers to deliver this information.

The day when, as political scientist Alan Rosenthal reports, a typical state legislator, asked about the legislative role in elementary and secondary education, replied quizzically, "Education is a local thing; we don't have anything to do with that; there's a formula", is over.[11] The schism between the technical and the political, which has led to research into the information needs and uses of professional educators while ignoring the political community, and the service agencies that support them, needs to be bridged.

It is the view of The Council of State Governments that the U.S. Department of Education -- acting uniformly or through an internal entity such as the National Center for Education Statistics -- should take the steps necessary to construct this bridge by planning and taking action to achieve three program goals:

-- Increase our knowledge and understanding of how state educational policy decision makers use statistical information to make decisions and about their information needs;

-- Assess the capacity of statistical information providers to assist state education policymakers to make better use of available statistical information resources, and offer statistical information providers insight into mechanisms for improving their services; and

-- Based upon this information user and provider analysis, develop a plan for improving the communication of useful statistical information to the non-SEA state educational policymakers.

When these goals are met, it is The Council's view that the U.S. Department of Education will have significantly added to our knowledge of how the process of governing education and making policy decisions might be made more effective. It will have also increased our understanding of the nature of program administration by keying on executive and legislative branch decision makers as information users and the national service agencies (such as The Council of State Governments) which support them.

It is the experience of The Council of State Governments that the provision of better information to state officials does result in better decision making. But information is a powerful tool only when it is provided to the right people, in the right way, at the right time.

The effort we suggest would assist the education community in forging more powerful information tools by providing the knowledge to give these tools better form and function.

The Information Needs of State Policymakers

In the early summer of 1978, a symposium was held in Arlington, Virginia, to review the experiences of various state and federal managers with the institutionalization of federal programs at the local level.

This symposium, sponsored by the MITRE Corporation, the National Institute of Education and the National Institute of Law Enforcement and Criminal Justice, focused on demonstration projects, but the comments made by the participants were telling from a number of perspectives. Mr. David L. Foote, Executive Director of the Colorado State Office of Planning and Budgeting, provided support for the effort we suggest when he told the federal program managers:

> If I had a single recommendation to make to the federal establishment ... it would be to take planning seriously. Not planning for demonstration, but what I would call policy planning, and making sure that we benefit from utilizing information that we continually generate and ask others to generate.[12]

As NCES has indicated through this requst for papers, the public policy challenges facing our system of federalism -- particularly in the area of education and its administration -- are formidable and can only be dealt with, as Mr. Foote indicates, by providing information to state policymakers in the most effective ways possible. But as we have noted above, these policymakers include a broad spectrum of public officials inside and outside of the SEA.

In studying how legislative, administrative and judicial policies and governmental organizations affect education, the most important questions may be: What statistical information is used by state policymakers to make decisions?; What information do they think they lack to make better decisions?; How do they wish the material to be presented to make it most useful?; What lessons can statistical information producers and providers learn from the information needs and wants of these state policymakers?

It seems, however, that because of the political/technical education schism, educational research has not focused on the information needs of decision makers in the political environment. One portion of the effort The Council would propose to NCES would be to attempt to study the information barriers to good education policy decision making which exist in that environment.

A project recently completed by the ERIC Clearinghouse on Teacher Education attempted to overcome one of the barriers to sound policy planning by identifying state policymaker information needs. But this study focused on three of the national state government service agencies rather than the policymakers themselves.

Several efforts have been made by the state government research community -- including efforts by The Council of State Governments -- to assess the information needed and used by state public policy decision makers. They have dealt, for the most part, with the elected members of the legislative branch, but have not considered conditions as they specifically affect education. These studies have centered on: the peculiarities of the legislature as an information-using institution; the kinds of general information legislators say they need; the interpretation of their needs by their staff; certain effects of the decision-making process on information use; and some constraints of the political environment on good information use. The findings of these non-education issue specific studies provide the research background for the first goal of the proposed effort.

In general, state officials (particularly state legislators) must make many decisions within a relatively short period of time. This is largely due to the fact that the volume of state business with which they must deal is rapidly increasing and the questions put before them often require quite specific and detailed knowledge of the issue area. As a result, the kind of information these officials need in order to make timely decisions is often not readily available.[13]

Moreover, researchers have found that these time constraints have forced the state official to become an information "schizophrenic." Generally the decision maker says that he or she wants information that is trustworthy, objective, reliable, comprehensive, applicable, and timely.[14] However, while they might be quite vocal about the quality and comprehensiveness of the information desired, they rarely seek elaborate information on policy issues. Indeed, in the state legislature this might lead to an information overload causing, "Paralysis (and) making things incomprehensible and unmanageable."[15]

As a result, although policymakers talk about their information needs, and on some topics their needs may be intense, they rarely seek elaborate information. When they get it unsolicited they do not know what to do with it.[16] Even legislative staff report a difficult time interpreting legislators' requests because of this duality of information needs.[17]

If the literature's portrayal of state legislators (and, although most of the research tends to deal only with the legislative branch, there is reason to believe that these conditions exist within the top-levels of the executive branch as well) is accurate, why should we be concerned with the statistical information needs of these actors? When we consider the "general" use of information, it does seem as though this schizophrenic information-seeking behavior would argue against any efforts to improve the provision of information relating to education issues.

But while this composite portrayal is essentially accurate, it is misleading when viewed apart from other aspects of the state decision making process.

Throughout a legislative session, for example, legislators will receive information they cannot use. This is because the information did not reach them at the right time in the legislative process; it did not help with problem solving; it did not support the decision makers' predispositions; it was not in a format which enabled the legislators to relate it to constituent needs; or it did not tell the legislators how colleagues felt about the issue.[18,19,20]

It is also a political reality that not every legislator or executive branch decision maker will be interested in every issue. Yet these decision makers will still face the prospect of having to make decisions on those issues either by endorsing them as executive policy or voting for or against them during the legislative session.

As a result, these officials will be forced to rely on decision making "shortcuts" in order to survive the flood of decisions that must be made. One method legislators have for making quick, but palatable decisions, is to rely on the orientations or predispositions they have brought to, or developed early in, their legislative careers. It is improbable that the individual legislator will have preset notions on every issue to be addressed, but it is likely that one or more of the legislator's colleagues will have some ideas about, or expertise in, a particular issue area. It is expected then--and supported by the research literature--that when an individual policymaker cannot arrive at a decision on the basis of personal judgment, he or she will look elsewhere for assistance. But to whom does he or she look?

Legislators look mainly to their colleagues and rely on their judgments.[21] A small group of individuals within the legislature is likely to be regarded as expert in a particular subject or issue area. Other members can usually rely on their ability to produce policies which reflect the values of the group as a whole.[22] Various studies have shown that policymakers tend to look within their own group for cues for decision making rather than to outsiders (such as SEA, federal education, or educational lab and center staff).[23,24,25,26] Lobbyists particularly have recognized the importance of seeking out those members of the policymaking group who are seen as opinion leaders.[27] Some education lobbyists have documented their strategies for informing these key legislators.[28]

These opinion leaders are also important in that they tend to occupy key positions in a two-step flow of communication between interested groups and individuals on the outside and the rest of their colleagues on the inside.[29] As a result, the ways in which these individuals with substantive knowledge make decisions, and the criteria they employ, will differ from that of their less knowledgeable, less interested, fellows.[30]

But it is unclear just how these non-SEA "opinion leaders" make their decisions, and what their statistical information needs are. More specifically, it is not clear how the non-SEA legislative and executive branch education opinion leaders make such decisions and employ statistical information.

On the basis of recent studies[31], we have some ideas as to why legislators become involved in education policy and are regarded as leaders in the area (e.g., legislator's background and/or district characteristics). On the basis of some isolated case studies, we also have some understanding of the factors which affect education policymaking in various states.[32] But these findings leave many questions unanswered and complicate our understanding of the decision making process as it relates to education:

--The literature suggests that education policy questions are handled differently than other policy questions by state legislatures, simply because of the nature of the political and educational environment.[33]

--The key opinion leaders in the field of education may not be make up solely of members of the education committees and education legislative staff, although that is usually where information studies related to education policymaking focus their attention. The leadership of the appropriations and finance committees, for instance, also have power over educational policymaking. A participant in education politics from a midwestern state notes, "They can have a nice time in house education talking about textbook selection, competency based education, and a lot of other things like that. Not much is going to happen on those things. It's the people who control the money who are calling the shots up and down the line."[34]

--The greatest source of information used by the state education community, the information produced by the federal government, is not widely used by the non-SEA public policymakers. Research on the influence of this information upon state political actors is surprisingly sparse. This is partly because it is relatively new, partly because the information is seen as being more useful to the school community, and partly because the federal information base on educational operations has not resulted in any theory-guided research literature.[35] Indeed, one study found that federal information providers constitute the last group state non-SEA problem solvers call upon for solutions.[36] Wirt has noted that the federal thrust for providing information to this group has been diffused, in part because, "Washington seeks to deal with complex organizations with a limited understanding of them."[37]

--The factor which may have the greatest effect on good information sharing between the education community and the public policy decision maker is the previously mentioned antagonism between the professional educator and the politician. Halperin notes that the schism between the technical and political actors continues with educators commonly saying that:

> "Politicans have a short term view of the world; their primary interest is in their own constituency and their narrow sectional, ethnic, regional or economic interests; and are poorly informed on educational issues." The politicians, on the other hand, say: "In order to frame social policy, we need facts, not generalities. We also need practical responses to immediate problems. Yet rarely do educators have the information we need to make sound policy; and, educators ought to know how to communicate, but there are few groups that speak less clearly, less concisely, and with more obfuscation. Instead of precise, comprehensible, here-and-now language, what we get is usually too olympian, too utopian, too abstract, or too fuzzy to be helpful."[38]

The Research Opportunity

What is clear, after a brief review of the literature concerning the use of and need for information by state public policy decision makers, is that the research community cannot now identify mechanisms for providing these political actors with more useful information on educational issues until more is known about the information environment.

As we have noted, studies of the use of information by the educational research community have historically centered on the policymakers in the state and local education agencies. Studies of the information environment of the state political actors have addressed general questions and have tended to favor studies of the legislature. These latter studies are even more limited--from an educational policymaking perspective--when we remember that the literature suggests that education policy questions are handled differently than other policy questions by state legislatures.

Additionally, recent studies in the education area (such as those by Rosenthal) have centered on only a portion of the political information network: the legislative education committee members and staff. The literature suggests that two-thirds of this network may have been overlooked: the decision makers who affect educational policy but are not normally seen as being part of the system (such as executive branch planning and budget officers and legislative branch finance committee chairpersons), and the education opinion leaders in both branches.

While informational barriers to better, more effective, state decision making exist in the political community, there is no evidence that these barriers are inherent to our political system. Barriers to providing adequate information for proper decision making were also noted in SEAs and LEAs, but have been greatly reduced by research efforts which have analyzed these barriers and suggested methods to overcome them.

What remains is the problem of identifying the best means of providing useful statistical information to the political decision makers within their unique environment. Given that information is "that which reduces error," we should conclude that providing better, more useable, statistical information to these important political actors will encourage improved educational policy decisions outside of the SEAs, in the same way that providing better information to SEA and LEA leaders has improved their capacity to make better decisions.

The research opportunity presented to NCES, then, is to: assess the information environment of the non-SEA education policymakers; review the results of this assessment; based upon this analysis, construct improved communication mechanisms to overcome the problems of legislative timing, relevance, personal predisposition, format, relationship to constituent needs and peer thinking; test these mechanisms; and disseminate successful approaches to the state educational policymaking community and those who serve them.

Conclusion

In its 1984 study of state response to the recommendations of the National Commission on Excellence in Education, The Council of State Governments found that state school leaders were taking the matter of reform for educational excellence to heart well before the Commission's report was finalized.[39] In this regard the Commission's work -- and the work of many other national taskforces assembled since the Commission's report was released -- serves as an additional propelling force for improvement efforts, but is probably not the initiator of the vast majority of state educational improvement efforts.

If America is at risk, and if the reforms outlined by the Commission are what are needed to achieve excellence, then there is every indication that the states are already taking the necessary actions.

But what was also clear from The Council's study is that state officials are attempting to look well beyond the Commission's report. Many states did not take the Commission's report for granted and numerous independent state and local task forces have been established to identify specific areas for improvement in each unique jurisdiction.

It is well that they have, for what seems clear is that for educational leaders to make additional headway in improvement efforts, additional experience is needed in communicating their "message" to the state political decision makers. The Commission gave weight to the importance of political actors at all levels of government, but is silent on the matter of how opportunities become programs and needs become budgets. There was every indication from The Council's study that the channels of communication -- and the quality and quantity of information carried by them -- which bring about informed educational opinion in state decision makers, need additional study.

The program goals The Council suggests to NCES in this paper would be an ambitious undertaking, but in electing to take action in this area The Council expects that the Center would be able to: accumulate extensive knowledge about statistical information as it is used by non-SEA executive and legislative policymakers in making decisions which effect education; assess the meaning of the findings as they relate to the information, dissemination and research programs of the educational research community, the federal education actors, state education agencies, local education agencies, and education public interest groups; and develop a model for improving the communication of statistical information to these actors.

The Council of State Governments would look with enthusiasm toward a long-range plan developed by NCES which would investigate the areas outlined above.

There may be no greater waste than information which goes unused. The Council looks forward to working with NCES to insure the widest and best use of statistical information by state decision makers.

REFERENCES

1. Cremin, Lawrence A. The Transformation of the School: Progressivism in American Education. Vintage Books: New York. 1964.

2. Wirt, Frederick M. and Michael W. Kirst. The Political Web of American Schools. Little, Brown: Boston. 1972. Pp. 5-11.

3. Wirt and Kirst (1972). Pp. 5-11.

   Katz, Ellis. the states rediscovered: educational policymaking in the 1970s, State Government (Vol. 53 - Winter, 1980). The Council of State Governments: Lexington, KY. P. 32.

4. Katz (1980).

5. Wirt, Frederick M. educational politics and policies, Politics in the American States: A Comparative Analysis, 3rd ed. (Herbert Jacob and Kenneth Vines, eds.). Little, Brown: Boston. 1976. P. 285.

6. Hines, Edward R. governors and educational policy-making, State Policy Making for the Public Schools: A Comparative Analysis. (Ronald Campbell and Tim Mazzoni, eds.). Education Governance Project, The Ohio State University: Columbus. 1974.

   Katz, Ellis. Educational Policymaking 1977-78: A Snapshot from the States. Institute for Educational Leadership, George Washington University: Washington, DC. Nov., 1978.

7. Katz (1980). P. 31.

8. Bank, Adrianne, et al. Dissemination and the Improvement of Practice: Cooperation and Support in the School Improvement Process. Far West Laboratory for Educational Research and Development: San Francisco. Feb., 1980.

   Cutter, Virginia. Dissemination Policies, Procedure and Programs of Nine State Education Agencies. Council of Chief State School Officers. 1974.

   Hood, Paul D. (ed.) New Perspectives on Planning, Management, and Evaluation in School Improvement. Far West Laboratory for Educational Research and Development: San Francisco. Nov., 1979.

   Hull, William and Ralph J. Kester. Perceived Effectiveness of Innovation Diffusion Tactics. The Center for Vocational Education. The Ohio State University: Columbus. Jan., 1975.

   Human Interaction Research Institute. Putting Knowledge to Use: A Distillation of the Literature Regarding Knowledge Transfer and Change. National Institute of Mental Health: Washington, DC. 1976.

Kester, Ralph J. and William L. Hull. Identification of the Empirical Dimensions of the Diffusion Process. The Center for Vocational Education, The Ohio State University: Columbus. Oct., 1973.

Louis, Karen, et al. Linking R & D With Schools. Abt Associates, Inc.: Cambridge, MA. Sept., 1979.

Madey, et al. Building Capacity for Improvement of Educational Practice: An Evaluation of NIE's State Dissemination Grants Program. Educational Research Center, NTS Research Corp.: Durham, NC. 1979.

Sieber, Sam D., et al. The Use of Educational Knowledge (Vol I-II). Bureau of Applied Social Research, Columbia University: New York. 1972.

Turnbull, Brenda J., et al. Promoting Change in Schools: A Diffusion Casebook. Far West Laboratory for Educational Research and Development: San Francisco, 1974.

Todd, William R., et al. A Review and Critical Analysis of the Literature Pertaining to the Diffusion of Educational Innovations. College of Communications, Florida State University: Tallahassee, FL. 1975.

9. Reed, Linda. Literature Synthesis: Marketing: Exploring Applications for Educational Dissemination. NTS Research Corp.: Durham, NC. 1978. Part II, P. 11.

10. Mattas, Frank and David Rawnsley. a justification for a user study. Unpublished manuscript prepared for a program sponsored by the Dissemination and Improvement of Practice Group, National Institute of Education. Summer, 1979. Pp. 18-19.

11. Rosenthal, Alan. Legislative Life: People, Process and Performance in the States. Harper and Row: New York. 1981. P. 271.

12. Chelimsky, Eleanor (ed.) Proceedings of a Symposium on the Institutionalization of Federal Programs at the Local Level. The MITRE Corp.: McLean, VA. 1979.

13. Porter, H. Owen. legislative information needs and staff resources in the american states, Legislative Staffing: A Comparative Perspective (James Heaphey and Alan Balutis, eds.). Halsted: New York. 1975. Pp. 39-59.

14. Rosenthal, Alan. legislative needs: what members want, what experts think, and what institutions can use, Legislative Priorities for Policy Assistance (Randy Huwa and Alan Rosenthal, eds.). Center for State Legislative Research and Service, Rutgers University: New Brunswick, NJ. 1976. P. 55.

15. Rosenthal (1981). P. 88.

16. Rosenthal (1981). P. 220.

17. Huwa, Randy, and Alan Rosenthal. Politicians and Professionals: Interaction Between Committee and Staff in State Legislatures. Center for State Legislative Research and Service, Rutgers University: New Brunswick, NJ. Jan., 1977. Pp. 18-19.

18. Rosenthal (1981). P. 222.

19. Huwa and Rosenthal (1977). Pp. 18-19.

20. Rosenthal (1981). P. 220.

21. Kingdon, John. Congressman's Voting Decisions. Harper and Row: New York. 1973.

Rosenthal (1981).

22. Mitchell, Douglas E. Social Science Impact on Legislative Decision-Making. National Institute of Education: Washington, DC. June, 1979. P. 202.

23. Menzel, Donald C., et al. Development of a Science and Technology Capability in State Legislatures. Center for the Study of Science Policy, Pennsylvania State University: University Park, PA. June, 1973.

24. Feller, Irwin, et al. Sources and Uses of Scientific and Technological Information in State Legislatures. Center for the Study of Science Policy, Pennsylvania State University: University Park, PA. June, 1975.

25. Castle, David S. State Legislative Information Systems: Development and Professionalization. Prepared for presentation to the Southern Legislative Conference, The Council of State Governments, Louisville, KY. July 17, 1979.

26. Rosenthal (1981).

27. Smith, Dorothy. In Our Own Interest: A Handbook for the Citizen Lobbyist in State Legislatures. Madrona Publishers: Seattle, WA. 1979. P. 4.

28. Porter, H. Owen. legislative experts and outsiders: the two-step flow of communication, Journal of Politics. (Vol. 36-Aug., 1974). P. 717.

29. Porter (1975). P. 52.

30. Rosenthal (1981). P. 88.

31. Rosenthal, Alan, and Susan Fuhrman. shaping state educational policy, The Interstate Compact for Education (Vol. 14-Fall, 1980). P. 22.

32. Bailey, Stephen K., et al. Schoolmen and Politics: A Study of State Aid to Education in the Northeast. Syracuse University Press: Syracuse, NY. 1962.

Masters, Nicholas A., et al. State Politics and the Public Schools. Alfred A. Knopf: New York. 1964.

Ziegler, Harmon, and Michael A. Baer. Lobbying: Interaction and Influence in American State Legislatures. Wadsworth: Belmont, CA. 1969.

Milstein, Mike M. and Robert E. Jennings. Educational Policymaking and the State Legislature: The New York Experience. Praeger: New York. 1973.

Iannaccone, Laurence, and Peter J. Cistone. The Politics of Education. ERIC Clearinghouse on Educational Management, University of Oregon: Eugene, OR. 1974.

Murphy, Jerome T. State Education Agencies and Discretionary Funds. D.C. Heath: Lexington, MA. 1974.

Patterson, Samuel C. american state legislatures and public policy, Politics in the American States: A Comparative Analysis, 3rd ed. (Herbert Jacob and Kenneth N. Vines, eds.). Little, Brown: Boston. Pp. 139-195.

Turnbull, Augustus B. III. staff impact on policy development in the florida legislature, Policy Studies Journal (Vol. 5-Summer, 1977). Pp. 450-454.

Katz (1978).

Katz (1980). Pp. 31-35.

33. Patterson (1976). Pp. 175-177.

34. Rosenthal (1981). P. 291.

35. Iannaccone and Cistone (1974).

36. Castle (1979).

37. Wirt (1976). P. 335.

38. Halperin, Samuel. politicians and educators: two world views, Federalism at the Crossroads: Improving Educational Policymaking. Institute for Educational Leadership, George Washington University: Washington, DC. 1976. Pp. 103.

39. Sims, E. Norman. The Recommendations of the National Commission on Excellence in Education: Reviewing the State Response. The Council of State Governments: Lexington, KY. 1984.

NEEDS FOR DATA IN EDUCATION

Considerations in Redesigning the Elementary and Secondary Data Program of the National Center for Education Statistics

William W. Turnbull
Educational Testing Service
Princeton, New Jersey

June 1985

NEEDS FOR DATA IN EDUCATION

Considerations in Redesigning the Elementary and Secondary Data Program of the National Center for Education Statistics

Reliable measures of educational status and trends are essential if we are to be able to monitor progress, understand change and set policy in education. That statement may seem a truism applicable to any aspect of the national interest but it is especially true in the educational realm. It is worth citing some of the reasons why this is the case.

The problem of deciding what data to collect is especially acute in education for several reasons. The first is the decentralization in this country of both the sources of information and the structures of decision-making in education. A second reason is that educational issues are intricately interwoven with and affected by a host of factors — economic, political, social, demographic — in the society, so that there are few natural boundaries to the relevant sources of information we may need to draw upon. Many non-educational agencies collect data that are related importantly to schooling but often they collect them in a form that precludes their easy incorporation into educational analyses.

Education's characteristics of dispersed control, varied record-keeping, diverse goals, and incompatible measures of results serve both to compound the data-gathering problem and to generate the tremendous need for a systematic, coordinated approach to a set of indicators of the health of the educational enterprise. Moreover, the diversity within the country offers the hope that reliable indicators, derived in comparable form from different parts of the country or from schools operating under contrasting circumstances, can pay big dividends in increased understanding of what practices seem to work well or poorly, so that we have a chance of not only describing our educational health but also improving it.

Among the key questions to be addressed are who will use the data for what policy purposes, and what information elements do they need to do the job? Further questions are what agency or agencies should collect the data, how should they collect it, and what role should NCES play in locating, assembling, reporting, and interpreting the information?

Audiences

There are multiple audiences for data about education. Since educational decisions are made at state and local levels, the information needs at those levels clearly must be met. The "local" decision-makers, however, need data not only about their own state or district but also about the nation. They need educational status and trend data on a multi-district, multi-state, national and international basis in order to compare their needs, efforts, and accomplishments with those of others.

Several groups must have national and international data because their responsibilities are national. They include federal legislators and administrators, research people examining the factors that influence educational progress or such questions as the interaction between education and economic development, and business people whose prospects for both manpower and markets are affected powerfully by what happens in education, both in the U.S. and abroad.

There is, then, a set of needs for data at the district, state, national and international levels that arise from the responsibilities of diverse groups of people. There is no way to distinguish the levels of aggregation of data needed by persons with broader or narrower geographic areas of responsibility: the broader picture provides an essential context for even tightly focussed local decisions.

Finally, the media and the public have a vital stake in the condition of education — a critically important "need to know." This need goes beyond raw data to a need for analysis and interpretation. This latter requirement is a hard one for any agency to meet in a way that will be perceived as even-handed but one that is nonetheless essential.

1. Need for A Program to Delineate Issues

The kinds of data needed obviously depend on the kinds of issues to be addressed. The delineation of the issues that the data should illuminate is a critical step and one that needs explicit attention. We need an intensive effort to develop a taxonomy of issues to be dealt with that will in turn generate the data to be collected. Second, we need a system for determining priorities within the taxonomic categories. Finally, we need to provide for regular review and modification of the taxonomy and priority matrices.

Collecting, analyzing, and reporting data is expensive, despite the powerful advantages offered by sampling of both respondents and data elements. There must always be trade-offs leading to the inclusion of some kinds of data and the exclusion of others, with the subsequent regrets when people find in perfect hindsight that some critical information is missing while information of other kinds is in oversupply.

While we will never be able to anticipate all questions, we might well be able to do a more satisfactory job of it if the effort to anticipate were itself made explicit and systematic. We need a taxonomy of issues within which to classify the questions to be asked, which in turn will generate a list of the data we need. The existence of an explicit matrix, giving shape and structure to the issues to be examined, would help to focus attention on the important policy questions at the most critical time, i.e. before the design for data collection has been decided upon.

A recommendation to NCES, then, is that a project be commissioned to develop a taxonomy of issues to be addressed by education data and related statistics. This project would be of greatest value to the present redesign effort if undertaken in the summer of 1985. Such a project could involve searching the literature in education and in other fields, preparing a set of discussion papers, and convening a working group of knowledgeable people to develop and publish a proposed taxonomy and priority system for widespread comment and for suggestions as to overlooked sources of data. The results would be of enormous value to the U.S. Department of Education in pursuing the redesign of its data program. The taxonomy would be subject to periodic revision and expansion as new questions are proposed and defined in future years.

An Interim Working Set of Issues

In the absence of the results of a specific effort to delineate issues, we may turn to useful statements now available to guide our thinking. An excellent example is the set of issues developed in 1983 as a guide to the reformulation of the National Assessment of Educational Progress.* Excerpts from the pertinent section appear below:

> **Policy Issues NAEP**
> **Should be Able to Address**
>
> It seems clear that NAEP must now serve a wide audience with diverse needs. Criticism of NAEP in the past has underscored its failure to be responsive to policy needs (Wirtz & Lapointe, 1982; Milrod, 1980; Wiley, 1981; Sebring & Boruch, 1982). What are some of the isues that NAEP should focus on as it reorganizes to meet the challenges of the eighties?
>
> Among the variety of pressing issues, three general policy areas stand out which should be addressed by NAEP because they require reliable data on student competencies and achievement: student competencies as they relate to national concerns; student achievement and attitudes as they relate to human resource needs; and, student achievement

---

*Messick, S., Beaton, A., & Lord, F. National Assessment of Educational Progress: A New Design for a New Era. Princeton, N.J.: Educational Testing Service, March, 1983, pp. 11-15.

as it relates to school effectiveness. In addressing these issues NAEP must not only be able to provide a national overview, but must also be relevant to state and local concerns — not for the purpose of needless comparisons among states or school districts but to assist individual states and localities in meeting their goals and objectives.

National Concerns

Since NAEP's inception, the federal government has designed and implemented education policies to provide equal educational opportunity to all citizens and to assure that young adults would be able to contribute to society in terms of both productivity and participation in the democratic process. The government clearly understands that an educated populace is a fundamental requirement for the nation's political and economic well-being. A major responsibility of NAEP should be to provide information for governmental and educational policymakers on the effects of their efforts and to act as an "early warning system" of potential problems. At a minimum, NAEP data should be relevant to the following kinds of questions:

> Are today's students learning the skills necessary for productive functioning in America in the 1980s? The 1990s? The year 2000?
>
> Are students in urban, suburban, and rural schools all being adequately prepared?
>
> Are public and private school children equally well prepared?
>
> Do children have access to programs preparing them to deal with the computer age?
>
> Are minority and disadvantaged youngsters being so prepared?
>
> What types of programs or allocations of resources seem to make a difference for disadvantaged and minority students?
>
> Are children from limited-English-speaking homes being provided the necessary skills?

Do students who have received special services under federal or state programs perform better than similar children who have not had access to those programs?

Do students leave formal education with positive attitudes toward productive work?

Human Resource Issues

The federal government is concerned with the flow of human resources to assure a work force competent to function in an advanced technology society and the necessary military personnel to protect American interests. Planning for human resource deployment is a complex process that requires reliable information on young people's competencies, training, and attitudes....

In the past we have vacillated between feast and famine in critical personnel areas.... NAEP should assist governmental and educational policy planners by contributing information on the following kinds of questions:

What are the career goals of high school students?

What are the attitudes of today's youth toward the military? toward business?

To what degree do students with access to science and high technology curricula choose careers in science more than those with no such experiences?

Are we preparing youth to meet the human resource needs in the health sciences? the humanities? teaching?

Are vocational/occupational programs equipping students with the skills they need to function in the work place?

School Effectiveness

School administrators are faced with rising costs and multiple demands on limited resources. They must choose among a host of competing interests. Achievement data, to be

> most useful, should be tied to other information to guide policymakers in deciding how they might best organize their programs and disperse their funds. Although achievement is influenced by many factors -- some school related, others beyond the school's control -- test data are one measure of the effectiveness of schools. Holding other variables constant, what factors within the purview of school administrators appear most likely to contribute to increased achievement? How can NAEP assist state and local policymakers to improve schooling?
>
> If NAEP is conceived not merely as a social indicator, but as a tool to identify problems and suggest areas of potentially productive research concerning educational progress, NAEP should attempt to provide data that address the following kinds of policy issues:
>
> How do pupil/teacher ratios appear to relate to achievement?
>
> Do students with preschool and/or kindergarten experiences seem to perform better than those without such programs?
>
> How do particular curricular approaches relate to student achievement in reading? writing? math?
>
> What are the relationships of in-service training programs, teacher turnover rates, and teacher competency requirements to student performance?

The NAEP-related list of questions is not sufficiently elaborated to serve the broad requirements of the NCES redesign. As an obvious example, it omits questions about the relative effectiveness of education in this country vis-a-vis others. A taxonomy of questions should include questions about how education is faring not only as compared with education in previous years but also as compared with education in other lands. Comparisons of student accomplishment in the United States with that elsewhere, accompanied by information about differences in educational practices associated with different results, can not only tell us how well we are "competing," it can also help us raise our sights in areas where others may be doing better and lead us to examine educational practices elsewhere that seem to be related to achieving better or worse results.

2. General Design Issues

A clear mapping of the questions to be answered from the data must be followed by a decision as to a strategy for gathering the statistics. The NCES activity is obviously not conducted in a vacuum. Many data-collection programs of other agencies, public and private, gather information that is directly or indirectly relevant to educational issues. In the single area of "educational outcomes," even a partial listing of large data bases that contain information derived from tests given in the U.S. contains over two dozen entries:

Precollege	Modal Age
National Assessment of Educational Progress (NAEP)	9
National Assessment of Educational Progress (NAEP)	13
Secondary School Admission Test	14
Metropolitan Achievement Tests	16

Iowa Tests of Educational Development 16

NCES 1980 High School & Beyond (HS&B) Sophomore Cohort 16

College Board Preliminary Scholastic Aptitude Test 17

NAEP 17

Armed Services Vocational Aptitude Battery 17

American College Testing Service 18

NCES 1980 Senior Cohort-Base Year Survey 18

Transcripts of High School Grade (from HS&B) 18

CEEB Admissions Testing Program (Scholastic Aptitude Test, Achievement Tests, Advanced Placement Examinations) 18

High School Equivalency Test 19

College and Beyond

Cooperative Institutional Research Program 19

NCES 1980 HS&B Senior Cohort--First Followup 20

NCES 1980 HS&B Sophomore Cohort--Second Followup 20

NCES Higher Education General Information Survey (HEGIS) 20

Graduate Record Examinations 22

National Teacher Examinations 22

Undergraduate Academic Transcripts (from HS&B) (17 to 22)

NCES 1980 HS&B—Senior Second Followup 22

NCES Recent College Graduate Survey 22

NRC Survey of Doctorate Recipients 26

Other surveys that are pertinent include those that are specialized by subject area (e.g. the RTI National Science Survey for ages 6-12), those that are international in scope but include the U.S. (e.g. the International Surveys of IAEA), and those that provide data only about other countries (e.g. the equivalent of our NAEP program, conducted in Great Britain). Still others that have developed large-scale longitudinal data bases over extended time periods, were sponsored by the U.S. Department of Labor in order to track educational and labor force activity. (These are found in the DOL National Longitudinal Surveys (NLS) of Labor Market experience and the Continuous Longitudinal Manpower Survey (CLMS).) Just developing (and maintaining) a good catalog of sources would help.

Trying to merge these data sets is a daunting challenge but an activity that should be studied and tried at least experimentally. A system of planned "linking sections" common to different data bases might prove feasible and helpful. Even if a complete merging is not feasible, some useful dimensions of comparability using subsets of data may be open to discovery and use. At least it should be possible in the course of such a study to develop good documentation, available centrally, about the comparability of the several files, including mundane but essential facts such as whether or not the data can all be run on the same computer! Such documentation would in itself be extremely useful.

The list includes examples of data from both governmental and non-governmental agencies. NCES already arranges to receive most of the pertinent data from the government agencies. The statistical series produced by non-governmental organizations in some cases are carefully maintained and are capable of illuminating special areas within the particular sphere of interest of the private organization. Cases in point are the data sets about college or graduate applicants, their test scores and their educational histories that are collected routinely by agencies like the American College Testing Program, the College Board and Educational Testing Service. The educational significance of these data in the public mind became dramatically evident in the 1970's when the persistent decline in the mean scores of SAT takers was first noted. NCES already draws on some of these sources of test data for their information on outcomes. A further step might be advance joint planning of issues that could be explored more effectively through cooperative arrangements similar to the agreements with SEA's or to existing interagency agreements within the Federal Government.

The thicket of problems becomes even thornier when one goes beyond data sets in a single broad domain -- educational outcomes -- and includes the many areas touched on in current population surveys by the Census Bureau and workforce surveys by the Department of Labor. Since the information in these data sets was not gathered on similar samples by asking a consistent set of questions, the job of NCES in trying to bring it together in relation to education issues is extremely difficult. The ideal (from one standpoint) of achieving complete comparability across data sets is impractical. The basic need to maintain continuity of long-running data sets is by itself a major deterrent to precipitate change. Nonetheless, efforts should obviously be made to remove unnecessary barriers to our ability to pool data across agencies. The present juncture, when NCES is in the redesign process, would seem to be a good time to explore the presently attainable degree of planned compatibility of efforts without jeopardizing the unique needs of each participating agency.

Frequently one finds that two surveys include questions intended to represent a whole complex of information describing a construct of common interest, such as socio-economic status, but have selected different questions. It would be worthwhile for an interagency authority such as OMB to study the extent to which the answers to different questions can be taken as valid surrogates for the broader construct.

Notwithstanding the serious obstacles to attaining compatibility across data files, we recommend that NCES take the lead in exploring with other agencies, public and private, the feasibility of achieving greater compatibility among data sets. An effort should be made to increase compatibility in the short term where possible or over a longer period where that is required. It may be that some highly desirable steps toward eliminating redundancy of effort that cannot be completed in the near future could be accomplished in 5, 10 or 15 years if started now.

3. "What" Studies and "Why" Studies

No matter how successfully NCES may be able to draw upon extant data bases to meet its needs, it will still need to make a very extensive primary collection effort of its own for several reasons: to fill in the gaps in some areas, to acquire data in a form compatible with its other data elements, or to fulfill its role as the principal data source in educational areas central to the purposes of NCES.

Many of the statistics collected (e.g. per pupil expenditure in public schools by state) are facts that have face validity as important in their own right. They answer legitimate questions of "what is."

The statistics on "what is" are more useful in answering questions about present conditions than in suggesting how or why those conditions came about. All too often, people juxtapose two or more sets of data about disparate conditions, find some instances of apparent correspondence, and infer a causal connection. The cross-sectional data of annual surveys are, of course, poor bases for causal inference. Much better for answering "why" questions are the kinds of data gathered in periodic studies such as the National Longitudinal Study of High School Seniors of 1972 (NLS) and the High School and Beyond Study of 1980 (HS&B).

The large scale longitudinal studies are proving to be critical in illuminating issues of public policy, since they provide a basis for tracing the later correlates of earlier student experiences, and they do so in the context of a wealth of background information. The background data — about financial support, interests, subjects studied, extra-curricular activities, and so on — help in interpreting the meaning of changes in attitude or in student learning or in decisions to continue or to drop out, or in changing job aspirations, both in general and differentially by such variables as sex or race.

We endorse strongly the view that the longitudinal studies are a uniquely valuable educational resource and urge that they be designed as a long-term and recurrent element in the NCES data-gathering system. Insofar as possible, regular long-term funding expectations (or, ideally, commitments) should be established and broadly announced so that other agencies could reaonably anticipate answering their questions on the basis of a continuing data series rather than feeling compelled to establish duplicative efforts.

Both the NLS and HS&B studies trace the progress of students from the high school years forward. In order to increase our understanding of what is happening in the pre-college years and why it is happening, we need a companion study that begins in the pre-school years and follows pupil progress through the grades, eventually linking up with the HS&B sample in secondary school. We urge that a "Preschool and Beyond" longitudinal study be instituted by NCES as early as possible. Such a study could be instituted as a stand-alone effort or possibly created as a longitudinal sub-study within NAEP. Ideally it should be undertaken on a broad national scale but if that approach is too expensive, thought could be given to the possibility of mounting it in a sample of cooperating states.

Data and Information

A legitimate question is how far NCES should go beyond gathering and reporting raw data by providing the analysis and interpretation that turn data into information.

In this country generally, the dramatic improvement of information processing technologies is likely to lead to a sharp increase in the volume of data recorded, manipulated and presented. We will have to be very deliberate in our procedures to avoid swamping audiences with undigested data "because it is there." The quality of synthesis and interpretation will have to keep pace if we are to realize the benefits of collecting the data in the first place.

It is our impression that at present the bulk of NCES's activity is devoted to providing data. It is our further impression that a growing component of NCES's work is in the areas of analysis and interpretation, through visual presentation and commentary in The Condition of Education, through the new publication on Indicators, and through a variety of special reports on particular topics. We applaud the shift in emphasis toward interpretation and encourage a continuation in the same direction.

We believe also that a strong effort should be made to encourage recognition of authorship of NCES's interpretive commentaries — another trend characteristic of recent years and one that could be carried still further. Signed analyses carry with them appropriate professional recognition for staff, with concomitant benefits in morale and career advancement. They also create at least some small theoretical distance between the responsibility attributable to the author versus the agency, even though in times of crisis that distance is usually very slight except in a purely academic institution.

General Issues of Strategy

Some of the other strategic issues to which NCES is no doubt giving attention in its review include:

> The periodicity of surveys and creation of public awareness of the schedule for collection of annual, biennial, decennial, etc. data
>
> Allocation of resources among cross-sectional and longitudinal studies
>
> Proportion of budget allocated to domestic and international surveys
>
> Reliance on NCES's own efforts versus dependency on data collected by others
>
> Extent to which data drawn from other agencies will be based on cooperative pre-planning versus serendipitous discovery
>
> Desirability of creating an interagency mechanism for coordinating data collection plans

The matter of how to encourage widespread use of the NCES data deserves intensive review. Possibilities such as more extensive use of networking to make the data readily available need continuous review as the available technology advances. Obstacles to, and techniques for encouraging, public use of data tapes should be explored, as should mechanisms for sharing insight and problems; e.g. creation of an AERA Special Interest Group for people using the High School and Beyond data sets, or creation of a consortium of data base users.

4. Specific Design Issues

The foregoing comments have been concerned with general issues of design. A more specific set of issues is posed if one asks a question like "what's wrong with the data we collect now? This is essentially the approach taken by Cooke, Ginsburg and Smith in their useful paper.*

Anyone who has worked with the NCES data or virtually any other data set will recognize and be able to add to the deficiencies reported in "The Sorry State of Education Statistics." The best way to improve the situation is a question of another order.

Many of the difficulties noted by Cooke, Ginsburg and Smith stem from NCES's dependence on the 50 states to collect and report data consistently. The cooperative arrangement with the SEA's has obvious advantages. Some agreement on common definitions seems essential, however, if the aggregated data are to be meaningful. It is suggested that NCES work through the Council of Chief State School Officers to procure comparable data from the SEA's. This need not disturb the individual state's internal definitions of variables such as attendance. A viable procedure might be to arrange for access to all of the raw data (e.g. number of enrolled children, number absent with excuses and without excuses, etc.) from which the SEA and NCES (or CCSSO for NCES) could derive statewide statistics to fit their own definitions. It would seem appropriate for NCES to stand ready to provide technical assistance to states that request consultation on the best ways of collecting and presenting their data.

---

*Cooke, C., Ginsburg, A., & Smith, M. "The Sorry State of Education Statistics," January 1985.

Some problems, like students' tendencies to overstate their course load, probably cannot be eliminated. It is suggested, however, that a continuing series of studies be undertaken of characteristic student response bias in key areas as a basis for deriving response adjustment coefficients. These coefficients could be applied systematically to provide more valid estimates of the true situation. The studies needed to obtain estimates of response bias would be intensive small-scale studies that would need to be repeated perhaps every 5 or 10 years, depending on the index. In some cases where the discrepancy between the response and the factual situation seemed extreme (e.g. where 80 percent of high school seniors report that they have taken a geometry course compared with 25 percent shown on transcripts) the most useful result would be clues as to how to revise the question rather than calculation of a response adjustment coefficient.

* * * * *

ETS is well aware that many of the suggestions made above may already be well represented in the procedures or the plans of NCES. We decided that in this paper we should err on the side of inclusiveness at the risk of redundancy. We will be glad to clarify points that need further explanation or to elaborate on ideas that may need exploration in detail.

THE LUTHERAN CHURCH–MISSOURI SYNOD

International Center
[illegible] South Kirkwood Road
Saint Louis, Missouri 63122-7295
314 965-9000 Telex 43-4452 Lutheran STL

June 11, 1985

Mr. Emerson J. Elliott
Administrator
United States Department of Education
National Center for Education Statistics
1200 19th St. NW
Washington, DC 20208

Dear Mr. Elliott:

The elementary and secondary schools sponsored by congregations of The Lutheran Church—Missouri Synod are interested in cooperating in the re-design of the education data program provided by the national center for education statistics. At this time we are not providing a formal paper for consideration in the first synthesis, but we wanted to be sure you were aware that we wish to join in the re-design and in the program itself.

It is helpful for us to know approximately how many non-public schools exist at what levels (preschool, elementary or secondary), how many children they serve, and how many teachers and administrators serve them. It is also helpful for us if we can separate The Lutheran Church—Missouri Synod schools from the other schools in your non-public school survey, and that we can compare their responses with those of the other church and non-church oriented private schools. It would also be helpful if we could compare data with the public schools.

Although we collect and disseminate rather extensive data on the schools of The Lutheran Church—Missouri Synod, receiving specific reports from over 95% of our schools, we are very interested in cooperating in this venture. We are eager to provide data for important surveys, such as the private school survey. In return we appreciate receiving the results of that survey so that we can continue to improve our schools.

It is important that the number of teachers in our schools be counted in the survey of teacher demand and shortage. Frequently the non-public school teacher demand and shortage is quite different than that found in the public schools. Information comparing both types of school would be appreciated.

One of the growing agencies in the schools sponsored by our church is extended daycare. This may become a service offered by public schools in the near future if federal financing should become available. I believe that information covering those schools which provide extended daycare before and after school would be important to be added to your statistics.

If I can help any further or if you have any questions about our data or data collection, please feel free to contact me.

Serving the Master Teacher,

Carl J. Moser

Carl J. Moser, Associate Secretary
Elementary and Secondary Schools

c.c. Dr. Vic Constien
Dr. James Boldt

DATA ON VOCATIONAL EDUCATION:
Problems and Recommendations

prepared for
National Center for Education Statistics
U.S. Department of Education

by

The National Center for Research in Vocational Education
The Ohio State University
1960 Kenny Road
Columbus, Ohio 43210-1090

June 20, 1085

INTRODUCTION

The Carl D. Perkins Vocational Education Act (P.L. 98-524) continues the information systems established by the 1976 amendments to the Vocational Education Act of 1963 (P.L. 94-482): a vocational education data system and the National Occupational Information Coordinating Committee. By these continuations Congress has reiterated its long-standing interest in better information both to assess the effects of the federal role in vocational education and to improve the working of the labor market. Any information systems established for vocational education must attempt to respond to these two objectives, but it is doubtful if any one system can do both.

The initial attempt to fashion such a system by the National Center for Education Statistics (NCES) failed because it tried to do too much. The original Vocational Education Data System (VEDS) attempted to provide information both for policy and for labor market purposes at a level of detail that local and state sources could not supply with acceptable accuracy. The internal and year-to-year inconsistencies in the data assembled by VEDS led the Office of Management and Budget to stop the collection of 1983-84 data.

NCES is currently trying to design a new "national vocational education data reporting and accounting system" which will comply with the mandates in the Perkins Act (Sec 421). These mandates are much the same as those in the 1976 amendments except there is a greater emphasis on special populations and on the use of sampling to collect data. This paper is intended to assist NCES in this process. It reflects the ideas of those staff members of the National Center for Research in Vocational Educationl who have worked most closely with the available national data on vocational education. First the problems that the original VEDS encountered are reviewed. The paper then presents recommendations for improving the operation of future systems and the utility of the data they collect.

Problems with the Old VEDS

A 1979 report by the National Center reviewed the early implementation of VEDS and concluded:

> VEDS future is still uncertain.
> . . . to the degree VEDS is implemented, vocational educators and decision-makers at every level will have a knowledge of who enrolls in vocational education programs, what happens to them afterward, and what it costs in the kind of detail long needed, but never before available. The extent to which it is implemented will depend upon the decisions of thousands of local and state administrators as they attempt to supply information in the form that VEDS requires (pp. 64-65)

By 1984 the uncertainty had been removed. VEDS as NCES had originally tried to implement it was stopped by OMB. VEDS had generated data at a level of detail not previously available, but the data were not consistent or credible. Comparisons of VEDS data to other sources yielded in a few states vocational enrollments that exceeded total secondary enrollment. Year-to-year changes within states in program enrollments of 50 to 100 percent were not uncommon. It would be easy to attribute the poor quality of these data to resistance or ineptitude among the data providers but more fundamental problems underlie most of these reporting difficulties.

Definitions

The primary difficulty at the secondary level is definitional. What criteria should be used to define vocational students? The quick answer is course enrollment: Students who take vocational courses are vocational students. By this definition, however, virtually all secondary students are vocational. (Campbell, Orth and Seitz (1981) have shown that over three-fourths (78 percent) of students take one or more courses designed to teach skills for paid employment. If consumer and homemaking and industrial arts courses are included, over 90 percent of students take at least one vocational course (Meyer 1981, NCES 1984).

Another technique frequently used is to ask students to classify themselves as to their main course of study. When comparison are made between self-report and other classifications made by administrators (Fetters 1975) or from an analysis of transcripts (Campbell, Orth and Seitz 1981) approximately one-third disagreement between the sources is found.

If additional criteria are applied to course taking data, such as total number of courses, areas of concentration, and the grade level at which the courses were taken, it is possible to distinguish those who appear to be preparing for entry into identifiable occupations from those who appear to be taking vocational courses for exploratory or avocational reasons. This is the approach that Campbell, Orth and Seitz (1981) followed and it enabled them to identify five discrete patterns of participation in vocational courses. These patterns distinguished between those who seem to be preparing for employment and others who took vocational courses without appearing to have employment as an objective.

The Campbell, et al. approach is applicable when students have completed or left high school and complete information is available on the courses they had taken. It is less appropriate for classifying students while they are still in high school. A student may, for example, take agricultural courses in the ninth and tenth grade, switch to auto mechanics in his eleventh grade and to carpentry in the twelfth grade. Depending on when the student was counted in his high school years, he would be classified in three different program areas. The Campbell et al. decision rules would classify such a student as a concentrator-explorer--one who initially appeared to follow an area of specialization but who left it to sample other occupational areas.

Reporting System

A second major problem the old VEDS encountered was the varied and decentralized nature of the system that generated the VEDS reports. The VEDS forms were distributed to the states. The information that was aggregated and reported on those forms was collected from local educational agencies by a variety of means. A few states virtually duplicated the VEDS forms and required the local agencies to complete them. Some states relied on individual student records which were completed at the local level and aggregated at the state level. Most states, however, tried to adopt their existing information systems to supply the information required by VEDS. The success of this approach varied widely across states.

Relying on such a decentralized system requires very good communication from the federal to the state and from the state to local levels. Even when the communication is good, there is an inevitable time lag between the initiation of a request at the federal level and the response at the local level. Any changes in the request, and there were many in VEDS brief history, compounds the communication difficulties. The repeated message that National

Center staff received when they contacted VEDS coordinators in the states was to stabilize and simplify the system. Of the two stability is probably more important than simplification.

Special Needs Populations

Vocational educators have had difficulty in developing accurate ways to identify disadvantaged and handicapped students for reporting purposes. These difficulties stem primarily from the discrepancy between the way special populations are served in schools and the way they are defined in legislation. The Perkins Act, for example, does not limit the definition of disadvantaged to income. It includes individuals "who have economic or academic disadvantages and who require special services and assistance in order to enable them to succeed in vocational education programs." [Sec 521(12)]. Even if the definition were limited to family income, public schools are reluctant to request such information. On those occasions when it has been requested, the schools have encountered resistance and protests. Consequently schools have had to rely on proxies of disadvantaged status, such as eligibility for free or reduced price school lunches. Even with such proxies, the names of those eligible are not widely shared and often are not available to the individuals responsible for completing the local forms that are aggregated for VEDS.

The existence of an individualized educational plan should be a clear indicator of whether a student is handicapped. Vocational administrators often claim, however, that many handicapped students are mainstreamed in vocational classes without the teachers or the administrators ever being informed of the students' handicaps. This is especially true among learning disabled and speech impaired students who constitute approximately two-thirds of all handicapped students. Their handicaps are less obvious and often less of a detriment in vocational classes. Furthermore, by the definition contained in the Perkins Act [Sec 521(15)], these students must require special education and supportive services to succeed in regular vocational classes to be considered handicapped. A literal interpretation of this definition means if they are not receiving special services, they should not be reported as handicapped.

Recommendations

This brief review of the major problems encountered by the old VEDS argues for a division between data collected primarily for policy purposes and data for labor market information. Most data for policy purposes can best be collected with special studies conducted on a sampling basis. Accurate data on program completers for the Occupational Information System, however, requires a census, a complete enumeration of the population of interest.

Data for Policy

Policy questions are basically concerned with who is served, how they are served, at what cost, and with what effects? Data to answer these questions can best be collected with specially designed questionnaires from representative samples of schools. The two on-going longitudinal studies of NCES and the one currently being planned can provide much of the needed data. Analyses of data of this type can provide far more precise information on the characteristics of vocational students, their secondary and postsecondary educational experiences and subsequent work careers than any aggregate reports.

The use of high school and postsecondary transcripts to define varying patterns of participation in vocational courses is recommended. Such a practice will deal with the definitional problem that plagued the old VEDS. The collection of original data from respondents in selected schools overcomes the difficulties of using the varied educational reporting systems in the separate states to generate the data.

Future longitudinal studies should be supplemented to provide more information on the educational process. Indicators of the educational process within vocational education could include:

- Student recruitment, selection
- Sources of curriculum
- Use of class time
- Relevance of equipment to that being used by employers
- Contact with business and industry
- Background of instructors, most recent experience in occupational areas they teach

Information for some of these indicators could come from the students as well as from teachers and administrators. Cost information can be obtained from the financial reports filed with the Office of Vocational and Adult Education.

Labor Market Information

Public vocational education is the major source of information on the supply of new workers for the Occupational Information System developed by the National Occupational Information Coordinating Committee. To provide this information at a level of detail sufficient for state and local planning decisions requires a census of all students who complete or leave vocational program after attaining a competency level judged suitable for paid employment in specified occupations. To attempt to provide these data on a sampling basis would require so many primary sampling units that the costs of collecting the data would be prohibitive.

It is recommended that instructors in public vocational programs complete a standardized form for each program completer or leaver. These instructors are in the best position to make such judgments on the competencies of their students. The form the instructors complete should contain background information on the sex, age, race/ethnicity of the students,[2] the program area in which trained and whether or not the individual attained a competency level suitable for employment in specified occupations. Such a report would deal with the problem of defining a program completer or leaver which is especially troublesome at the postsecondary level. Most postsecondary students do not enter a specified program. They take one or two selected courses to meet personal needs. Many of them are employed at the time they take these courses.

The forms instructors complete would be submitted at periodic intervals by local educational agencies to their state offices. At the state level, the forms would be aggregated for use in the state occupational information system and a cumulative report made each year to NCES. The National and State Occupational Information Coordinating Committees would like program completers to be reported at the six digit level according to the Classification of Instruction Program (Malitz 1981) code. This level does not reflect the way most vocational programs at the secondary and postsecondary level are offered. These programs are designed to provide preparation for employment in a number of related occupations. To require reporting at the six-digit level forces individuals to make arbitrary choices that cause unreliability in the data.

The six-digit level may be appropriate for some short-term retraining or upgrading courses, but for most longer-term vocational programs it is too specific.

The total number of programs completers reported by public vocational programs should not be entered directly as data on the supply of skilled workers. Campbell, Gardner and Winterstein (1984) have found that less than half of secondary students who complete extended vocational programs actually seek employment immediately after high school. Many go on for additional post-secondary education. The results of this and future research should be used to adjust the supply data for the Occupational Information System so that the completer figures more accurately reflect those who actually seek employment.

FOOTNOTES

1. In subsequent discussion National Center shall refer to the vocational center and NCES to the statistical center.

2. These data would be for policy not labor market purposes.

REFERENCES

Campbell, Paul B., Gardner, John A., and Winterstein, Paul. Transition Patterns Between Education and Work. Columbus: National Center for Research in Vocational Education, The Ohio State University, 1984.

Campbell, Paul B., Orth, Mollie N., and Seitz, Patricia. Patterns of Participation in Secondary Vocational Education. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1981.

Fetters, William B. National Longitudinal Study of the High School Class of 1972: Student Questionnaire and Test Data by Sex, High School Program, Ethnic Catogory and Father's Education. Washington, D.C.: National Center for Education Statistics, 1975.

Malitz, Gerald S. A Classification of Instructional Programs. Washington, D.C.: National Center for Education Statistics, 1981.

Meyer, Robert H. An Economic Analysis of High School Vocational Education I. Vocational Education How Should it be Measured? Washington: The Urban Institute, 1981.

National Center for Education Statistics. "Patterns of Participation in Secondary Vocational Education." NCES Bulletin (84-213b) August, 1984.

National Center for Research in Vocational Education, The Status of Vocational Education: School Year 1976-77. Columbus: The National Center for Research in Vocational Education, The Ohio State University, 1979.

Suggestions For The NCES Redesign Project

A Paper Submitted By

The
National Education Association

June 15, 1985

Suggestions For The NCES Redesign Project

The National Education Association (NEA) is entering the information age with a lack of information. In a nation accustomed to numbers, this statement seems absurd. After all, public and private collections of educational data exist. Statistics bombard the education market weekly. The public press turns out bales of educational reports yearly.

All of this information should provide a clear picture of the status of education in this country. It should provide sharp reflections of the way education is changing. It should also help answer with increasing sophistication the many questions of cost, benefit, and quality. Available information, however, does not reflect well the educational landscape.

Bringing education into sharp focus is as difficult today as it was twenty years ago. The need for a sharp focus, however, is perhaps greater today than it was in the past. Government regulation of education has increased. State funding for education has increased. The number of students enrolled in schools is increasing. And new measures to reform education are everywhere present.

All of these changes carry with them a demand for more and better information about education. Yet the current supply of data has not kept up with the demand, and educators have reason to worry. If the demand is not met, then future policy will likely be based on a murky picture. Furthermore, we will not know with any certainty what impact efforts to reform education have had.

The redesign project of the National Center for Education Statistics (NCES) offers a partial but significant solution to the information problem. The Center's resources, capabilities, and professional reputation are ideally suited for the data-collection efforts needed in education. Although the NEA does not believe NCES can or should solve all the information problems, we do believe it can solve some of them. For this reason, the NCES redesign project is of major importance, and the NEA is pleased to be involved.

NEA suggestions for the NCES redesign of its elementary and secondary education data program are organized below under five categories suggested in NCES guidelines. The categories are: Issues and Data Needs; Data Modifications; Importance of NCES Series; Data Deletion; and Relevance, Quality, and Utility.

Issues and Data Needs

Today, our national self-perceptions of education are regularly confirmed or challenged by statistics on many matters. Whether the meanings read into the data are reasonable or fanciful, the numbers provide a basis for popular and specialized discussion.

The NEA expects official numbers, especially those that appear in series, to play an increasingly prominent role in policy deliberations. Among the many issues likely to be discussed in the future, several noted below seem particularly amenable to NCES collection efforts.

Effective Schools

Issues of quality education will likely expand to include recent effective schools research. Subjects of interest can be expected to include characteristics of school organization, school governance, school administration, public expectations for schools, federal and state regulations, and local policies.

The current NCES data program will not provide sufficient support for the effective schools issue. The following revisions should help strengthen the data base:

- Add to both the public and private school surveys data elements pertaining to the school characteristics of organization, governance, administration, expectations, regulations, and policies. Suggestions for such elements appear in Tables 1 and 2 at the conclusion of this paper.
- Consider expanding the NCES program to include case studies, field studies, policy reviews, historical research, and additional surveys to expand the scope and detail of effective schools data.

Equity

Since World War II, numerous policy changes concern the issue of equity. School busing (to adjust the numbers of white and black students) and job quotas (to ensure the efficacy of affirmative action) are but two examples.

Equity issues pertaining to race and sex will likely persist. Issues pertaining to age and ethnic origin will likely grow. The following suggestions anticipate the demand for more detailed data where they are not already gathered:

- Refine the variants of Spanish ethnicity to include Mexican American, Puerto Rican, and Cuban for all NCES surveys.
- Refine the variants of Asian ethnicity to include Pacific Islanders, Japanese, Chinese, and Vietnamese for all NCES surveys. Consider using the U.S. Census item for this refinement.
- Add the ethnic category of American Indian and Alaskan Native for all NCES surveys.

Public and Private Schools

A broadened perspective on schooling is evident in current policy discussions. This perspective includes K-12 schools in the mixed public-private system of schooling, pre-primary schools, and adult education and training programs in both public and private sectors.

This perspective raises many questions about public and private schools: how they are similar, how they are different, and what they can learn from each other. The following redesign suggestions anticipate a continued and broadened public-private school debate:

- Add to the Sample Surveys component a Pre-Primary School Survey designed to gather information about the location, organization, program, governance, finance, employees, and students of these schools.
- Align the data elements in the Private School Survey and Public School Survey so that the two surveys are comparable. Specific suggestions for revisions appear in Tables 1 and 2.
- Expand the scope of data collected for both the private and public school surveys. Suggestions for expansion appear in Table 2 concluding this paper.
- Add to the Sample Survey component a survey or to the Other Agency Data component survey items that track the magnitude and growth of adult education and training programs in both public and private sectors.

School Finance

Limited resources require that money be spent on education wisely at the local, state, and federal levels. Current revenue and expenditure data seriously curtail the kinds of questions that can be asked and answers that can be explored. Because the demand for cost-effectiveness studies will likely increase in a period of limited resources and fiscal restraint, the need for more detailed revenue and expenditure data will grow. The following suggestions for NCES revision can ease the need for better fiscal data:

- Provide greater revenue and expenditure detail for both private and public schools. See Tables 1 and 2 for specific suggestions.
- Add data for incentive plans and salaries of public and private school administrators and educational support personnel.

Student Outcomes

Statistics are regularly published on such fundamental matters as reading and literacy rates, achievement rates, and dropout rates. The inadequacy of measures for each of these rates is well established and well known. The following suggestions call for a major revision of student outcome data:

- Expand and standardize the definition of student performance outcomes to include more knowledge areas. Consider such categories as linguistic, musical, logical-mathematical, spatial, bodily-kinesthetic, and personal knowledge.
- Develop measures appropriate for an enlarged view of student outcomes. Consider the possibility of building upon the diversified measures developed for the National Assessment of Educational Progress.
- Collect fall and spring enrollment figures.
- Standardize definitions of dropout, attendance, and literacy.
- Convene an advisory group to study methods suitable for measuring dropout rates and student mobility.

Teaching Quality

For a number of years, fairly simple models directed the collection of data pertaining to teaching quality. For example, some models were constructed on the metaphor of an assembly line with students and procedures standardized. Some models viewed students as passive agents over which teachers had absolute control. Other models, including those underlying several reform proposals, take an economic view of teachers and teaching.

Accumulated research and the experiences of teachers indicate that all of these models misrepresent reality. For this reason, the issue of teaching quality will likely become more complex as models for thinking about quality change. Suggestions consistent with this change include:

- Add data elements for each of the following known components of effective teaching: personal characteristics of teachers, teacher competence, teacher performance, student learning experience, student learning outcomes, teacher credentials, school context, characteristics of students as a class, and characteristics of students as individual learners. Suggestions for data elements appear in Tables 1 and 2.
- Consider convening an advisory panel to develop and refine over time measures of teaching effectiveness.
- Consider expanding the NCES program to include case studies, field studies, policy reviews, historical research, and additional surveys to expand the scope and detail of data for the components noted above.

Data Modifications

NCES data bases should be expanded to provide a broader and more detailed source of information about K-12 education. Specific recommendations for expansion appear in Tables 1 and 2. Table 1 identifies data elements that need to be added to existing data collection components. Table 2 identifies data elements that should be collected through additional surveys. Additional surveys are needed to provide a more comprehensive assessment of the setting, characteristics, practices, personnel, conditions, and outcomes of K-12 education at the local level. Suggested data elements in the table would be useful in providing estimates of:

- A wider array of local conditions related to school, teacher, and pupil performance.
- Variation within schools, districts, regions, and the nation in the presence of conditions likely to influence education quality.
- Many dimensions of education practice at a given time and place.

A major desirable outcome of expanding the NCES program will be an improved data base for professional, governmental, and public decisions regarding ways to improve the quality of K-12 education. Such data can contribute to a better understanding of the complexity, variation, and similarities of K-12 education throughout the nation. The data can be used to study problems, opportunities, and decision options related to improving education. The data can also be used to study departures from traditional approaches to funding, accrediting, evaluating, and changing public schools.

The utility of NCES data bases can also be increased if the following criteria are consistently met:

- Conditions in education are accurately measured and reported.
- Data are collected with instruments that meet high technical standards.
- Data are collected under normal conditions.
- Data are made available to users in a timely fashion.

Importance of NCES Series

Statistical information produced by NCES shapes thousands of decisions by government and nongovernment users alike. The need for these data has escalated during the eighties as researchers, legislators, and the public focus on issues of educational improvement.

The importance of NCES' statistical programs has grown in conjunction with this expanding need. Consumers of education data have come to view NCES data series--all series--as accurate, non-biased sources of information with which to address an increasingly complex education enterprise. While each NCES data series has a wide audience of users and may be considered as essential to the planning and design of public education policies, the Common Core of Data may represent the most heavily used series of public school statistics.

The Core is the cornerstone of educational information in the United States. No other public or private institution collects and maintains public education data to the extent that NCES does via the Core. Several groups--including the NEA--conduct data collection activities which parallel the Core in some respects, but these efforts pale in comparison to the NCES program. For this reason and others, the support and maintenance of the Core component should be a national priority.

The Core represents the most basic data series within the NCES. It enables assessments of what was, what is, and what will be in a statistical sense. Annual updates to Core surveys provide basic statistical information on public schools, their pupils, personnel, and finances.

This information, by school, LEA, or state, serves as an invaluable tool for measuring change, signaling trends, and designing future education policies. Data from the Core are used as benchmarks, goals or objectives, and measures of success and failure. Simply stated, the Core allows us to "know things" about public education, and thereby allows us to "do things" about public education.

The Core stands ready to support basic research, budgeting decisions, and programmatic planning. It can uncover questions and issues requiring further investigation. Data collected through the Core, then, are of immense value in and of themselves, but they also serve as means to the achievement of numerous goals.

Educational issues of concern to government officials, educators, and the public throughout the remainder of the twentieth century will call for education information as provided through the Core. Indeed, selected variables from the Core can and have been used to help assess the outcomes of recent reform initiatives in various states. This evaluative process is vital to the prospects for meaningful reform of public education offerings.

Data Deletion

The NEA recommends against any reductions or elimination of existing NCES data series. We believe that all the series and items are vital to the interests of the educational community.

Relevance, Quality, and Utility

The summary documents and data tapes provided by the NCES have been useful to the NEA. NCES' access to school and school district data and their data collection processes cannot be matched by any nongovernmental organization. NEA believes that the work of NCES is vital to our own efforts in service to the education community. NEA does have specific recommendations for improving the data, the collection process, and the dissemination of the data in order to increase the relevance, technical quality, and utility of the data programs.

The relevance can be improved by providing more timely data, access to more raw data, consistency between public and private school surveys, and consistency in surveys over years. The speed with which the data are made available to the public is critical. The availability of the raw data, on tape, permits NEA and others to perform their own analyses. The use of the same questions for public and private school surveys permits more extensive matching and comparing of school systems on a wide variety of attributes. The same issue of consistency applies to surveys repeated over the years. The use of the same questions permits an analysis of trends.

The technical quality can be improved by ensuring complete and accurate documentation, more complete editing of the data, and increased efforts to eliminate missing data. Data provided on tape should be thoroughly reviewed for errors in record descriptions and data documentation. More comprehensive edit checking would reveal inconsistencies in the data. For example, the computation of ratios between certain items such as enrollment and teachers would highlight unreasonable data configurations that do not appear in individual items. The use of random audits for individual schools and districts may reveal ambiguous data.

or areas of difficult data gathering. For example, a review of individual districts may reveal that the definitions of staff categories used by districts do not match those of NCES. Efforts to revise or to promote use of those codes could then be undertaken. The elimination of missing data could provide more accurate summary data.

The usefulness of the data can be improved by providing more documentation on the availability of NCES data and more contact with NCES personnel for future survey planning. NEA needs to know what data are available from NCES, in what forms the data are available, and when the data are released. Increased contact between NCES and the user community will enhance the use of present data and the planning of future surveys. The process for future survey planning that is now being implemented is an excellent step and should be maintained.

Table 1. Summary of NEA Suggestions for Additions and Changes for NCES Data, By Census and Survey

NCES Component	NEA Suggestions
Common Core of Data	
1. Public School Universe	o Add spring membership. o Add full-time-equivalent classroom teacher by sex and elementary/secondary level.
2. Local Education Agency Universe	o No additions or changes.
3. Local Education Agency Nonfiscal Report	o Add fall membership by grade. o Add number of full-time-equivalent LEA employees in all employee categories. o Add number of full-time-equivalent teachers by individual grade. o Add presence or absence of collective bargaining agreements for teacher, administrator, and educational support personnel groups.
4. Public School District Finance Report	o Provide revenue by source consistent with NCES handbook on financial accounting. o Provide expenditure by function consistent with NCES handbook on financial accounting. o Provide other uses of funds by category consistent with NCES handbook on financial accounting. o Provide special exhibits by category consistent with NCES handbook on financial accounting.
5. State Aggregate Nonfiscal Report	o Add fall membership by individual grade.

Table 1 continued

NCES Component	NEA Suggestions
Nonfiscal continued	o Add full-time-equivalent employees by major assignment category, by state.
	o Add number of high school graduates.
6. State Aggregate Fiscal Report	o Make revenue, expenditure, other uses, and special exhibits detail consistent with revisions suggested for district finance data.
	o Add average daily attendance.
	o Add state law defining average daily attendance.
	o Add state aid formulae.
Sample Surveys	
1. Private School Survey	o Add fall membership by individual grade.
	o Add total membership.
	o Add design capacity of school.
	o Add ethnicity enrollment as percent of total enrollment.
	o Add grade span.
	o Add teacher college credits by subject matter field.
	o Add information on additional training for teachers.
	o Add personal characteristics of teachers.
	o Add membership by major subject.
	o Add average SAT/ACT scores and percent tested by school.
	o Add teacher incentive plans.
	o Add teaching assignment and classroom enrollment for teachers.

Table 1 continued

NCES Component	NEA Suggestions
Private School continued	o Add teacher hours per week by activity.
	o Add number of volunteers by activity category.
2. Public School Survey	o Add fall membership by individual grade.
	o Add highest degree earned by teacher.
	o Add number of years experience by teacher.
	o Add admission requirements, disciplinary policies, length of day and school year, and other characteristics consistent with private school survey.
3. Recent College Graduates Survey	o Develop better descriptors for this data set.
	o Add SAT/ACT scores.
	o Add academic program/preparation detail.
	o Add more demographic detail.
4. Survey of Teacher Demand and Shortage	o Provide rationale for number of private schools in survey sample.
	o Add number of teachers leaving and why.
	o Add number of full-time-equivalent teachers by grade.
	o Add class size by type of class.
	o Add number of budgeted positions.
	o Add number of needed positions.
	o Add descriptions of recruitment and employment practices.
	o Provide rationale and greater detail for teacher incentive plans.

Table 1 continued

NCES Component	NEA Suggestions
5. High School And Beyond	o Add descriptions of teacher and student recruitment and placement policies.
	o Add more male guardian questions on the parent questionnaire--whether he was present during pre-school years, worked, etc.
	o Add community descriptive elements such as racial/ethnic mix, community size.
6. Library/Media Center Survey	o Add number of books, materials loaned.
	o Add number of computers, programs available.
Other Agency Data	
1. Preprimary Enrollments of Children 3 - 5 Years Old	o If this survey is repeated, add length of program day, program year.
	o Add enrollment and placement requirements for students.
	o Add health care services descriptions.
	o Add state law governing provision of preprimary education.
	o Add state law governing attendance.
	o Add more program description.

Table 2. Recommended Data Elements By Major Category For New Or Expanded NCES Sample Survey Component

---

A. Students
   1. Fall membership by grade
   2. Enrollment by subject matter
   3. Students qualified for special programs
   4. Students enrolled in special programs
   5. Average SAT/ACT scores and percent of students tested
   6. Ethnicity status and percent of total enrollment
   7. Student transfers
   8. Student dropouts
   9. Student attendance
   10. Experience with violence
   11. SES distribution

B. Teachers
   1. FTE by school level
   2. FTE by category (regular, special education, etc.)
   3. Assignment type (department head, chair)
   4. Tenure status
   5. Job differentiation status (Master teacher, mentor teacher, etc.)
   6. Activities in day
   7. Highest earned degree
   8. Years of experience
   9. Education specialty in college
   10. Salary average per school
   11. Salary intervals per school
   12. Additional training
   13. Credits by subject, continuing education
   14. Marital status
   15. Sex
   16. Age intervals
   17. Race

C. Administrators
   1. Highest degree earned
   2. Years experience
   3. Education specialty in college
   4. Salary average per school
   5. Marital status
   6. Sex
   7. Age intervals
   8. Ethnicity
   9. Administrator salary schedule
   10. Administrator salary by intervals

Table 2 continued

D. Educational Support Personnel
  1. Standard definitions and classifications
  2. FTE by job category
  3. Compensation plan, fringe benefits
  4. Job qualifications by category
  5. Staff development practices
  6. Evaluation practices
  7. Supervisory practices
  8. Demographic data

E. Finance
  1. Revenue by source (include private and federal grants)
  2. Expenditures by major assignment category: compensation and fringe benefits
  3. Expenditures by major classification

F. Programs and Practices
  1. Length of school day
  2. Length of school year
  3. Programs offered
  4. Class size by class type
  5. Pupil load
  6. Admission policies by type of school
  7. Teacher activities by percent of time
  8. Teacher incentive plans
  9. Teacher education programs
  10. Teacher support programs
  11. Teacher evaluation process
  12. Discipline policies
  13. Job differentiation plans
  14. Standardized testing programs
  15. Grading policies
  16. Span of teacher authority

G. Classroom
  1. Classroom size
  2. Classroom space
  3. Books and materials
  4. Audio-visual equipment, high tech equipment
  5. Available supplies

H. School Administration
  1. Source of authority (public, private: religious, nonreligious)
  2. Structure (single school, state, private network)
  3. Location of school (city, suburb, rural)

NGA COMMENTS ON NCES
REDESIGN OF THE ELEMENTARY/SECONDARY
EDUCATION DATA PROGRAM

Introduction

A majority of the governors indicated in their 1984 and 1985 state of the state messages that education continues to be a top priority. The general theme of the 1984 and 1985 state initiatives has been to improve the quality of the education system. The governor as the state's chief executive officer guides education policy in primarily three ways. First, the governor performs a leadership role in setting the agenda for and promoting activities on educational improvements/reform. Second, the governor is responsible for developing state budget requests for education expenditures. In many states the education budget accounts for as much as 30 percent of the total state budget, and third, many governors are responsible for appointing state education board members.

As leaders in setting the state education agenda, governors in 1985 planned to focus primarily on initiatives to improve the teaching profession, address student quality through improvements in math and science instruction, and examine and increase the financing of education. This represents a shift to more specific strategies from the broader initiatives of the past few years to link educational reform and improvements with other state policies directed at promoting state revitalization and economic growth. Table 1 indicates that in 1985 the largest number of states expressed an interest in the issues concerning incentives in the teaching profession. Twenty three states emphasized teacher career ladders and 21 states emphasized teacher salaries. Of the remaining top ten issues, 4 are related to school administration, 3 of these to financing mechanisms and 1 related to school administration reform. Another four issues are related to student quality including student competency and child abuse. In comparison, in 1984 the governors in 29 states emphasized their interest in the broader aspects of building stronger partnerships between education and business/industry. Private sector linkages were seen as a means to accomplish the goals of preparing a better-educated future work force and thereby contributing to the state's future economic growth potential.

Table 1

Top Ten Education Issues Cited in the 1985 State of the State Messages

Issue Area	Number of States Citing
Teacher Career Ladder and Development	23
Teacher Salaries	21
Math/Science Instruction	20
Finance Formula	20
Economic Development, Voc. Ed., and Tech.	19
School Administration Quality and Reform	18
Funding Increase	18
Local School Aid	18
Student Testing and Competency	17
Child Abuse	17

Source: Governors' State of the State Messages, 1985 (Forty-seven state messages reviewed.)

During the 1980's, more than 30 states have enacted major financial reforms of elementary/secondary education. In 1984, 25 governors planned to enhance education efforts by providing more fiscal resources and 9 governors proposed tax increases to provide adequate resources for the improvement of their states' educational system. The majority of the funding increases proposed (20 states) was planned for use in raising teachers' salaries while 9 states proposed funding for improvements of school facilities and equipment. Ten states proposed increasing state aid to local schools systems. In 1985, 20 governors were interested in finance formula issues, 18 governors plan to enhance education efforts by providing more fiscal resources and 18 governors emphasized local school aid. This represents a significant shift from the 1970's when more than 28 states enacted major financial reforms for elementary and secondary education to equalize fiscal resources among school districts in order to relieve property taxes and legal challenges.

State Policy-Makers Data Needs

In order to perform education policy setting functions, states need to plan, develop, implement and evaluate education initiatives. For these purposes a combination of state and national data are useful. State produced education data which are designed to meet specific state needs provide the core for state education policy-making. All state education policymakers rely heavily on information from the State Education Department but also on local school districts and on education associations. However, national trend data and consistent and accurate data from all states for macro comparison purposes is of key interest as well.

The top ten priority issues as identified in the governors' state of the state messages can be used to provide a framework for a discussion of data needs. The list of the top ten education topics is provided to give a picture of the current education topics of interest to governors. We are not suggesting, however, that in trying to determine the needs of the data system for elementary/secondary education for the next 10 years, that currently popular issues such as those identified in Table 1 be used exclusively as a barometer of long-term data needs. This list, however, does reflect back to some broader core issues that fluctuate little over the long run.

The education citations in the state messages can be organized into three general issue areas which include improving student, teacher and school administrative quality. Table 2 illustrates this organization and ranks the topics in each of the three areas. As can be seen from this table, the primary topics of interest to improve student quality concern basic education focusing on technology and technical skills. Other areas include competency measurement, child care and well being, student groups at the extremes (dropouts and gifted) and special-need students; and community behavior that affects student outcome. The issue of graduation requirements ranks last. The key issue in improving school administration is school finance. Other issues include general management initiatives. Of somewhat lesser interest is the issue of teacher shortage.

The list of issues in Table 2 can be examined in terms of more specific data items to determine which are of interest to state education policymakers. The Education Policy Consortium developed a preliminary list of potential data items of interest to consortium member associations and their constituencies. These data items listed in Table 3 are organized for this paper into four categories including student data, teacher data, school data and finance data. The data items related to the governors' top ten priority issues are indicated with an asterisk as are the items generally available through NCES.

In the first area in Table 3, student data, the governors are not timid to talk about measurement of educational outcomes across states. As more states move toward preparing a better educated workforce to encourage economic development the issue of identifying student outcomes emerges as more than assessing student achievement. More data than test scores, such as the SAT, are needed to determine post-school experiences. The education process should be traced from start to outcome to determine what happens to the in-school population upon leaving an education program (by graduating or dropping out), what are their post-school labor market experiences in terms of employment, unemployment and earnings, and whether they re-enter school at some future time. In the future, student outcome measures may be one set of evidence used in evaluating education reform policies currently being initiated.

Longitudinal studies such as the High School and Beyond Survey are one way of determining outcome measures. This is one of the few surveys that capture data from students on a longitudinal basis. It seems, because of its somewhat unique nature that this questionnaire should be a priority to be maintained, improved in terms of data quality and potentially be expanded to gather more data, in terms of content and sample size to make the data more state specific.

In the second area, outlined in Table 3, teacher data, the governors are currently interested in examining the teaching profession as a primary factor in improving the education system. Incentives to keep and attract quality teachers,

Table 2

Summary of Education Initiatives

Cited in the Governors' 1985 State of the State Messages

A. Topics Cited to Improve the Quality of K-12 Students

Rank	Topic	Number of States
1	Math & Science Instruction	20
2	Economic Development, Voc. Ed. & Tech.	19
3.5	Testing and Competency	17
3.5	Child Abuse	17
5	Child Care & Early Childhood	15
6	Dropouts & Discipline	13
7	Gifted	12
8	Computer Literacy	11
9	Community & Parent Involvement	10
10	Special Ed. & Handicapped	8
11	Graduation Requirements	4

B. Topics Cited to Improve the Quality of K-12 Teachers

Rank	Topic	Number of States
1	Career Ladder & Development	23
2	Salaries	21
3	Certification & Evaluation	11

C. Topics Cited to Improve Quality of K-12 School Administration

Rank	Topic	Number of States
1	Finance Formula	20
3	Administrative Quality & Reform	18
3	Funding Increase	18
3	Local School Aid	18
5	Management/Class Size	13
6	Facility/Equipment/Text Books	11
7	Teacher Shortage	4

Source: Governors' State of the State Messages, 1985. (Forty-seven state messages were reviewed.)

Table 3
Preliminary List of Education Policy Consortium
Identified Data Items Related to the Top Ten Governors' Priorities

Topic	Data Item Description	Governors' Top Ten Priorities	NCES Data Available
STUDENT DATA	Achievement	*	*
	test scores		
	school grades		
	promotion record		
	Attainment	*	
	drop out rates		
	post drop out experience		
	graduation rates		*
	post graduation experience		
	In-School Behavior		
	attendance/truancy		
	vandalism		
	suspension/expulsion		
	course enrollments	*	*
	attitudes		
	Community Behavior		
	voter registration and participation		
	Individual Characteristics	*	*
	demographic (age/race/sex)		
	SES background		
	grade level		
	type of school attending		
	migrant/refugee		
	primary language		
	handicap		
	abused as a child	*	
TEACHER DATA	Training/Certification		
	formal education		
	participation in:		
	in-service training		
	pre-service training		
	loan/scholarship availability		
	certification subjects		
	Evaluation of performance		
	classroom evaluation		
	competency test scores		

Table 3 (Continued)

Topic	Data Item Description	Governors' Top Ten Priorities	NCES Data Available
	Compensation	*	*
	pay for performance		
	incentives		
	salaries/benefits		
	Employment Status		
	retention rates		
	reasons for leaving profession		
	working conditions	*	
	years of service		
	course assignments	*	*
	Personal Characteristics		
	demographic (age/race/sex)		
	attitudes		
	academic talents by employed/ leaver		
SCHOOL CHARACTERISTICS DATA	Curriculum and Assessment Improvement	*	
	Instructional materials		
	Training		
	Alignment of curriculum within grade, across grade, with assessment		
	LEA and state testing		
	Education indicators other than tests		
	School Improvement	*	
	strategies		
	training		
	planning process		
	curriculum improvement		
	State Role		
	monitoring/accountability		
	technical assistance		
	relationship to local districts		
	Technology		
	Community Involvement/Satisfaction with Schools		

Table 3 (Continued)

Topic	Data Item Description	Governors' Top Ten Priorities	NCES Data Available
FINANCE DATA	State and Local Revenues	*	*
	tax base level and compo.		
	tax rate level and compo.		
	total revenues		
	sources of revenues		
	relief provisions		
	revenue limitation		
	provisions for non-public		
	Expenditures (local)	*	*
	total expenditures		
	total elementary/secondary expenditures		
	expenditures by category		
	cost of special programs		
	student aid categorical vs. formula aid for special student population expenditure limits		
	Federal Aid	*	*
	total amount by state		
	allocation formulas used (federal and state)		
	state/local split		
	administrative services split		
	source of match		
	$ amount of audit except		
	$ amount of carryover		

Source: This preliminary data item list was derived from an Education Policy Consortium meeting, February, 1985.

such as compensation and career ladders are of key interest. Of particular concern is the notion of higher salaries generating more qualified teachers, which if validated, would encourage states to change salary structures, increase the minimum and/or average teacher salaries and establish merit or incentive pay programs. Of lesser interest to governors is the number of teachers employed and teacher shortages. If the profession has attractive incentives, the issue of teacher shortage may be of negligable importance.

The Public and Private School Surveys which both collect information on these topics should be among NCES priorities. The samples should be examined to determine the feasibility of expansion to collect data more state specific. This should be considered in conjunction with the further examination of appropriate state administrative records and a deemphasis on the Teacher Demand and Surplus Survey.

In the fourth area of Table 3, information on school finance has been an issue of interest for a number of years and will no doubt continue to be of interest as efforts to improve the school system are continued into the next decade. While the issue is not a new one, a shifting focus overtime to different aspects of school finance is evident. The equity issue of the 1970's in funding local school districts has shifted to interest in funding formulas, budget increases and increasing teacher salaries. Of interest in the future will be financing issues of public versus private schools. Basic finance information concerning both sectors should continue to be considered a core data element in any elementary/secondary education data system. Particularly in the Private School Survey, attention should be given to the finance questions. To enhance their data base, NCES should also examine the finance data base maintained by ECS.

Data Duplication/Overlap

If the NCES data collections, as presented in Attachment B of the initial correspondence concerning the redesign, are examined by subject area rather than by data collection program some duplication of effort becomes evident. For example, if the topic of information on teachers is examined then it appears from a cursory review that several surveys collect information on teachers in addition to the data available from administrative records. We recommend a thorough review of the Survey of Teacher Demand and Shortage, Recent College Graduates Survey and the Public School Survey to examine duplication. All three ask questions of teachers concerning subject matter assignments. The Public School Survey and the Survey of Teacher Demand and Shortage both ask questions concerning teacher incentive plans. The Recent College Graduate Survey as well as the Public and Private School surveys obtain teacher salary and compensation data.

In addition, there is duplication between NCES and other federal agency collections. For example, the Bureau of Labor Statistics, as part of the Occupational Employment Statistics (OES) program collects data every three years on current teacher employment. Another form of duplication is when information is collected in a survey but is already available in administrative records. Information on the number of new hires and the number of those teachers returning to their previous position from the survey of Teacher Demand and Shortage could potentially be obtained through the Employer's Quarterly Wage and Tax Report of State Employment Security Agencies (SESA).

We would recommend that because of the apparent duplication the Survey of Teacher Demand and Surplus be given a very low priority and reviewed to determine if it should be conducted at all. It seems that the data could be obtained primarily from the BLS OES program, state administrative records and if needed through the Public and Private School Survey. As well, the recent College Graduates Survey should be examined to determine what information is received that can't be derived from College/University and SESA administrative records and/or the Public and Private School Surveys.

Although the area of teacher data seems to have the largest number of separate collections and therefore the greatest potential for duplication, other areas as well may also have inefficiencies. Beyond this one example cited above which needs further examination, we suggest that NCES do a comprehensive review of their data collections across subject areas to explore further efficiencies that could be realized through unduplicated data collection and more extensive use of administrative records. A single collection instrument that obtains relevant data for multiple purposes and users appears to be a far more efficient use of resources than multiple shorter surveys resulting in several sets of incompatable data.

Unmet Data Needs

A data system to remain relevant to users should be flexible in meeting data needs created by the changing nature of our society. Although several unmet data needs can be identified, no priorities have been assigned to these needs by state policy officals. For example, the High School and Beyond Survey traces the post-high school experience of graduates but not of those youth who are not graduating or who have dropped out. This will become an increasingly important topic to determine the experiences of the at-risk population and the impacts of the policy initiatives of the early 1980s. Again, administrative records may be a tool useful in gathering some of this data or it may be necessary to explore collection of these data through the High School and Beyond Survey.

States need to identify education outcomes related to their own labor markets to fully use the data in their own policy development process, because of different industry structures, different rates of growth/decline and different labor market barriers and characteristics. For this reason, states do not find useful national longitudinal education data that only report national estimates, estimates for the 9 census regions and the 7 largest states. Although 6 or 7 states paid to have data collected for a state specific sample, generally there is no state level detail available.

We acknowledge funding limitations and because of this are supportive of NCES efforts that would be more creative in developing arrangements to have states expand their samples. Also NCES should consider expanding the national sample to provide more state specific data.

Another gap is information on career ladders within the teaching profession. Although there is a general lack of this type of information for most occupations, with the emphasis being placed on incentives for teachers, this information is valuable. Methods to obtain this data should be considered by NCES working cooperatively with the National Occupational Information Coordinating Committee.

Issues That Crosscut Specific Data Programs

The NCES efforts to assure quality data is produced as recommended in the NCES "Research in Statistics" report are critically important. Governors and state policy representatives are seriously concerned about the accuracy of reported data. It is generally acknowledged that there are few validity studies made or audits done on NCES data collections. We therefore strongly support work to develop procedures and strategies for continually assessing the validity of all NCES data programs and encourage NCES to make this a top priority. There are other efforts that would also improve the quality of data; such as developing minimum definitions that would make data consistent across states.

Standardization of Data Needed

NCES data are most useful to state policy makers when comparisons can be made between states and between a state and the national trend or average. For example, the indicators of education status and trends are an excellent vehicle from which comparions can be made if the defintions used across states are consistent. Some states even indicated to us that it would be preferrable to have a less extensive data collection effort that contained more rigidly defined data elements that were released more quickly.

States are very interested in assessing the amount and method of teacher compensation. In a survey of Governors' education policy staff conducted in November 1984 by the State Education Policy Consortium 43 percent of the states responding in the area of teacher quality indicated that the single issue which would be the most important over the next year is teacher compensation and is likely to remain so for the next decade.

In this case, state specific information on the state education budget, current compensation level and structure, the state history of salary increases by local education agencies, comparable personnel costs within the state, the elements included in the compensation package, and teacher characteristics may be used. National information on the current national median teacher salary, whether it is rising/falling and by how much, the variance of each state from the median and where each state ranks compared to others gives the state a relative measure of their teacher salary program against a national indicator and neighboring states. Salary information should be collected in a consistent manner.

Comparative measures that are not based on standardized definitions to guide data collection and that do not include a description of what's reported and what's not reported in the data may lead industry and education policy analysts to erroneous conclusions. For example, consider the detailed comparisons of data across states for the retirement system. In some states, local government pays the employer contribution and in others the state government pays the employer contribution. A state could have relatively higher or lower figures based on who pays for the retirement system and how the data are collected. Another example is enrollment data which, depending on the state, could represent average daily enrollment or a head count.

Numerous other examples exist of data collections where no standard definitions are used and data are not compatible across federal data collectors. An example is the long term debate on how to define a teacher. The BLS/Census and NCES definitions both differ. While we do acknowledge that there are different uses for different types of data, because of the need to integrate and use a variety of data sources in answering policy questions discrepencies such as this make it extremely difficult to accurately interpret the data being examined. Instead of using different definitions in various data collection programs we believe that it is the responsibility of the federal government to use common definitions for data collection.

To overcome the definition inconsistencies such as these that render the data meaningless for the key purpose to which they are used at the state level will require a sustained long-term effort to determine what are the core elements and how they should be collected. While education systems do vary widely across states; it appears that states would welcome common reporting on certain national data elements that would allow valid comparisons to be made.

Fifteen years ago BLS had similar problems with the definition of unemployment. NCES may wish to look at the BLS federal/state cooperative programs
as an example of how national definitions are used particularly with state administrative records. The BLS defines those data elements which are necessary for the national income accounts data. The BLS then contracts with the state to collect data using these standard definitions; using the dollars as a leverage tool.

The BLS as the major statistical agency responsible for labor force statistics has defined the population (16 years and older) into mutually exclusive categories as shown in Figure 1. NCES as the major statistical agency responsible for education statistics, should consider defining the population (0-16 years old) in a similar fashion of mutually exclusive categories. This would help in the development of definitions.

Standardizing data is not done without problems. BLS has been taken to court over the definitions but has won all the cases. Definitional problems should be carefully examined through more extensive collaborative efforts between the federal and state levels. A useful mechanism may be an interagency approach which brings together users and producers such as that of the National Occupational Information Coordinating Committee.

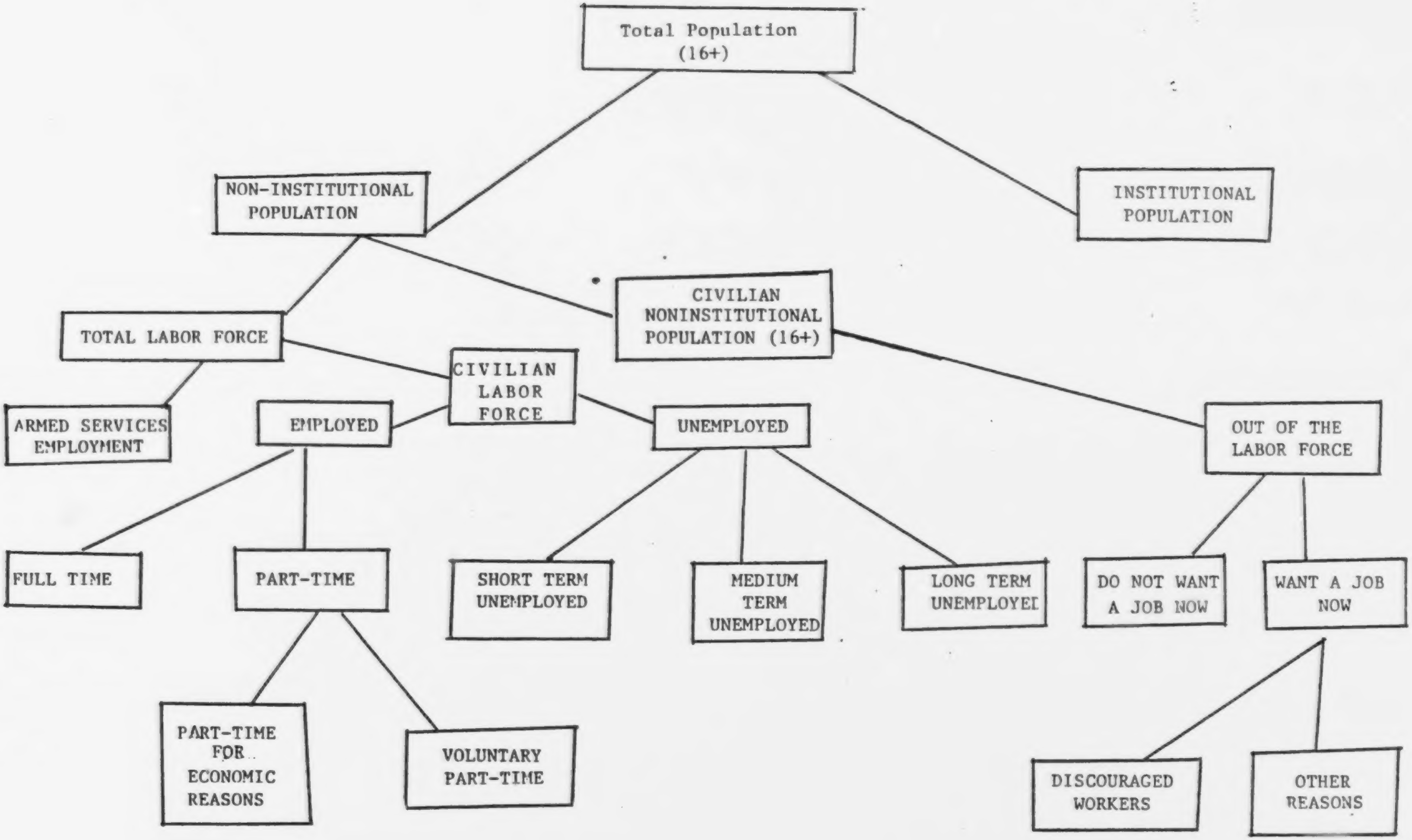
Statistical Agency Coordination

Under the authority of the Federal Paperwork Reduction Act we would recommend you pay particular attention to coordinating education financial data collected through the various programs with the Census Bureau which collects state and local fiscal data through the Census of Governments. The Census Bureau could potentially collect the data for NCES under an interagency agreement or contractual arrangement collection similar to the CPS October education supplement collection.

Other areas where interagency coordination would be appropriate are with the Department of Labor's Employment and Training Administration. Longitudinal

Figure 1

Labor Force Concepts
Used in the
Current Population Survey


Total Population (16+)
NON-INSTITUTIONAL POPULATION
INSTITUTIONAL POPULATION
CIVILIAN NONINSTITUTIONAL POPULATION (16+)
TOTAL LABOR FORCE
CIVILIAN LABOR FORCE
ARMED SERVICES EMPLOYMENT
EMPLOYED
UNEMPLOYED
OUT OF THE LABOR FORCE
FULL TIME
PART-TIME
SHORT TERM UNEMPLOYED
MEDIUM TERM UNEMPLOYED
LONG TERM UNEMPLOYEI
DO NOT WANT A JOB NOW
WANT A JOB NOW
PART-TIME FOR ECONOMIC REASONS
VOLUNTARY PART-TIME
DISCOURAGED WORKERS
OTHER REASONS

survey data from a variety of sources including the JTPA training programs would be more useful if coordinated among agencies and made more compatible. Under the JTPA, a longitudinal job training survey is conducted using a national sample of individuals representative of youth, all adults and welfare adults conducted on an annual basis. Data are collected on the individuals at 12 months and 24 months after program completion. We would suggest here some sharing of information between NCES and the Employment and Training Administration. There may be questions of interest to NCES which could be added to the JTPA Survey and as well NCES could add questions of interest to ETA to their survey. The coordination of these sources of data would provide an expanded data base for more extensive use and would be a more efficient use of limited resources.

Release Raw Data Quickly

In many situations data are needed quickly for policy purposes. For example, if states want to change their salary level and/or structure in relation to the current national average trend, then data that are several years old are not useful. The early data systems being recommended by the NCES "Research in Statistics" report is one step to address the issue of timeliness. However, states will have limited uses for these data due to the small national sample sizes. State education policy makers would like quicker turn around time on the administrative data submitted to NCES directly by the states. An emphasis on technological initiatives that improve data collection, editing and processing procedures and data release capabilities will aid in shorter turn-around times.

It has been suggested by several state data users that the administrative data collected from the states be made available in raw form to provide earlier access by state analysts. Use of NCES data may well increase if the turnaround time from the point when states report data to the point when NCES disseminates the collective data sent back to states is shortened.

It was suggested by several states that NCES should concentrate less on detailed analysis of state-specific data or inter-state comparisons and more on national trends and a look to future issues. State experts who are familiar with the particular qualities of the area being analyzed and the differences between states should be primarily responsible for this analysis. Interpretation of national data should be conducted based on a set of standards which maintain the integrity of the data.

Conclusion

There is general support for NCES data products and publications in the states although this is not the primary source of data used for state education policy-making. The states as partners with the federal government are committed to the reporting of state/local data under reasonable requirements to various federal agencies such as NCES and national organizations. In some states this reporting activity may require up to 3 to 4 person-months per year. In exploring the return on the state's investment in NCES data and the usefulness of the nationally collected state education data several key issues stand out from a state policy perspective. By far, the most important issues cross cutting a majority of states are that of:

(1) improving the quality of data;

(2) providing comparability of data across states, across data programs, and across data collectors;

(3) the increased use of state administrative records such as those available through the SESA; and

(4) capitalizing on the statistical expertise in other federal government agencies and associations.

NSBA

NATIONAL SCHOOL BOARDS ASSOCIATION
1680 Duke Street, Alexandria, Virginia 22314
(703) 838-NSBA

June 27, 1985

Mr. Emerson Elliott
Administrator
National Center for Education Statistics
1200 19th Street, NW
Washington, DC 20208-1631

Dear Mr. Elliott:

We are pleased to comment upon the Elementary/Secondary Education Data Program proposed by the National Center for Educational Statistics (NCES). The redesign project is a welcome effort on your part, not only because it is commendable of NCES to consider the need to update its efforts but also (and especially) because we appreciate the dedication and thoroughness with which NCES is seeking counsel and advice from throughout the community of educational statistics users. We hope that this NCES process will serve as an exemplar to other offices of the Department of Education.

The most useful service to local school district policy makers, and to the state school boards associations that assist these local policy makers, is that federal education statistics capture significant trends in a timely way. The current effort is imperfect, in this regard; it catches some but not others. For example, the movements toward magnet schools and toward desegregation are not readily traced in NCES's data, while changes in the nature of the teaching profession are more readily found there.

The next decade -- because of the "excellence movement," the changing demographics of the U.S. population, and the emerging electronic technologies that many hope will improve society at large and schools in particular -- is likely to be an era of change in public education. Capturing trends, in a timely way, will require some reshaping of NCES's inquiries.

In light of this general conceptual background, we respectfully offer the following comments and recommendations:

- We anticipate a continued policy interest in improving the effects of the schooling enterprise. We need to shift our statistical focus more than we have, to assist the public debate about school effectiveness. Certainly the public's interest reflected in the "excellence movement" has been framed in terms of increasing student learning -- not only academic learning measured by traditional tests but also other kinds of learning: The changing nature of employment has prompted attention to thinking skills, computer capability, job-seeking and job-holding skills. Higher education institutions have expressed concern about levels of writing and study skills. Some of the nation's social ills have focused attention on a gamut of values held by students and graduates, ranging from entrepreneurship to patriotism and racial tolerance to sexual responsibility. The focus on effects of schools should reflect this broad range of learning.

The current NCES program, however, generally focuses its attention on what goes into the schooling process. This historic imbalance has an unintended consequence: the sheer weight of data reported about the number and characteristics of teachers, courses taught, attendance, enrollment categories, family characteristics, et al., leaves the impression that education is more interested in counting our resources than counting our effects. Moreover, the emphasis on statistics about the schooling process assures that process trends will more likely be captured than trends in the effects of education, so that the public debate gets frustrated. We hope that NCES through this redesign project will find new ways to aid the public's interest in the effects of schooling.

- New electronic technologies, and new systems for management and instruction that exploit these new tools, will serve and reinforce this public emphasis on the effects of schooling, in two ways:
    - The arguments for and against uses of computers and other technology will be formulated on the basis that new tools do (or do not) improve the product of the schooling enterprise; and
    - Some of these new tools enhance the ability of school management to gather timely data about the success of the schooling enterprise.

We need data about the uses and impact of technology. There is much talk and some considerable action in the uses of new technology but little useful data. Only market sales data is currently available routinely. The current ad hoc study by the Research Triangle funded by NCES on Computers for Instruction in Higher Education deserves a parallel study for elementary and secondary education. (Some of the questions in that study, that focus on policies and academic requirements suggest some ways to address some of the issues about the effects of schooling also.) It would be very useful to know not only what devices schools are buying, but also to what uses these devices are being put, how staff is being trained, how courses and budgets are being altered et al.

Another aspect of technology's impact will be changes in the roles of school professionals. Current federal statistics seem to suggest that all teachers are alike; and that they work in classrooms. We expect to see further differentiation among school staff roles, fueled not only by pressures for career ladders and merit pay but also by the introduction of technologies. Already, for example, the "computer teacher" that serves in a "laboratory" as a resource to many other teachers has a very different role from the conventional image of what teachers do. Other new professional roles and titles and circumstances are likely to emerge as schools use television, computers, electronic mail, on-line data bases and laser-disc-based data bases and other new technologies.

Two other dimensions of the introduction of new technologies can be anticipated: more diverse instructional strategies and more diverse student roles, as the new mediated and individualized tools become commonplace.

- The National Assessment of Education Progress (NAEP) is a vital federally-financed program because it provides a continuing measure of some effects of our schooling system. This, and perhaps other "snapshots" by NCES can give a picture of changing student attitudes (that is, attitudes that schools may have a role in shaping) to supplement data about academic learning.

- Four topics about the schooling process are not well captured in the current data program that should be improved in the next decade, as they are likely to be of policy concern:

    - courses available to students;
    - populations served by schools;
    - off-campus learning; and
    - uses of new technologies.

We recommend special attention to these four topics.

New administrative procedures, and both old and new technologies, make it feasible to broaden the menu of courses that can be taught even in the smallest schools. On the other hand, pressures to get "back to basics" and to shift academic priorities tend to reduce the menu of courses available in larger schools. What to teach, and how to make it available, are constant policy topics at local and state levels. Trend data would be helpful here.

The federal statistics program seems to define public schools in K-12 terms. Yet schools everywhere are looking at pre-school care, after-school care, and many forms of adult education and services. Whom to serve, and how to render services to new populations, are policy issues at local and state levels which NCES may be able to illuminate with trend data.

The issue of on-campus versus off-campus learning suggests several dimensions. Technology makes home-based learning more feasible and school-building-based learning less necessary. One trend is the growing interest by some school districts in uses of broadcast, public and cable television. Another is the often expressed need for school/business collaboration; related is a potential for greater collaboration between schools, libraries, museums and higher education institutions.

In summary, schools are likely to change in several ways during the next decade; the best NCES service would be to capture significant trends in a timely way.

We see a most significant trend in the shift in the public debate from debating the process of education to debating its effects. NCES data should also make the shift to facilitate this trend.

We hope that these comments and recommendations will be helpful to you as you participate in the laudable effort to improve NCES data for practical use in local school districts throughout America.

Very truly yours,

Thomas A. Shannon
Executive Director

TAS/mk

cc: William J. Bennett
Secretary of Education

RESEARCH AND DATA NEEDS FOR SMALL/RURAL SCHOOLS

A Position Paper prepared for the
National Center for Education Statistics

In consideration of NCES's plans for the
Redesign of their Elementary and Secondary
Data Program

by

*Bruce Barker
Box 4110
Texas Tech University
Lubbock, Texas 79409

June 17, 1985

*Bruce Barker is an Assistant Professor of Education and Assistant Director of the National Center for Smaller Schools at Texas Tech University. He also serves on the Research Committee of the national Rural Education Association. This article has been prepared at the request of the Rural Education Association.

RESEARCH AND DATA NEEDS FOR SMALL/RURAL SCHOOLS

by

Bruce Barker

Although discussion of research and development needs for small/rural schools is becoming more prevalent, the state of knowledge and information currently available on rural education remains incomplete and startlingly inadequate. The first National Seminar on Rural Education held in Washington, D. C. in May, 1979 recognized the need to collect, analyze, and compile data pertinent to rural education (Flectcher, 1979/80). Tamblyn (1977) indicated that one of the major tasks in the 1980's for rural education was the need to conduct basic research on small school problems, practices, and unique features. Horn (1981) declared that one of the responsibilities facing universities is to conduct research and collect data on rural schools. And, Nachtigal (1979) specifically stated that descriptive data are needed on the operation of K-12 rural school systems with enrollments of fewer than 300 students, 300-999 students, and 1000-2500 students.

The Problem of Definitions for Rural Education

Lack of a precise definition may be one reason rural education has received little attention in recent years. Rural education has been a difficult entity to define because the word "rural" has different meanings when viewed historically, statistically, or philosophically (Salmon, 1980). Furthermore, the concept of "rural education" varies from state to state and region to region. Both Texas and Oregon, for example, define a small/rural district as having fewer than 1000 students (Barker, 1985). Since 1970, the U.S. Census Bureau has carefully defined the rural population as consisting of all persons living in places of fewer than 2500 inhabitants or in areas of extended cities with a population density of less than 1000 persons per square mile (U.S. Department of Commerce, 1970). The National Advisory Council on the Education for Disadvantaged Children uses the Census Bureau's definition to state that a rural district is, therefore, one having fewer that 2500 students (Special Report on Rural Education, 1979). Other definitions include those of the American Association of School Administrators which has established a K-12 enrollment of 2000 or less as a small district and the National Association of Secondary School Principals which considers an enrollment of 1000 or less to be a small high school (Jinks, 1984).

Some of these definitions imply that rural America collectively consists of all our society's nonmetropolitan areas. Inference is also made that this portion of our society is basically homogeneous in nature and composition. In reality, rural America is a vast array of diverse nonmetropolitan areas which may be internally more homogeneous that most urban communities, but which differ widely from each other. For example, an island hamlet off the coast of Maine, an Alaskan native village near the Arctic Circle, a coal mining town in West Virginia, a ranching area in Wyoming, an impoverished community in the Mississippi Delta, a ski resort section of Vermont, or a prosperous grain farming region in Iowa have little in common, except that they are all classified by the Census Bureau as rural areas of the United States (Sher, 1977).

The Lack of Attention Given Rural Schools

Federal statistics reveal that 59.5 million Americans live outside designated urban areas of the United States and that rural school students constitute the largest minority public school population in this country (Treadway, 1984; Sher, 1977). Based on the Census Bureau's definition of "rural," nearly two-thirds of the 15,600 operating public school districts located in the United States are in rural areas and one student in every three attends an elementary or secondary school classified as rural (REA News, 1982). Ironically, however, the "lion's share" of attention, research and an over balance of federal and state financial support generally go to large schools in metropolitan areas. Not until late 1983, four years after the establishment of the U. S. Department of Education, did that federal department declare a "Rural Education and Rural Family Education Policy for the 1980's" which stated, "Rural education shall receive an equitable share of the information, services, assistance, and funds available from and through the Department of Education and its programs" (ERIC CRESS, 1983/84). In 1983, the National Center for Education Statistics also agreed, for the first time, to include small and rural schools of under 300 students as a separate category for data collection (REA News, 1983). Up until the time of these two actions, national policy makers and researchers had paid little attention to rural schools.

Rural Education Research and Data Needs

Among the expressed goals of the national Rural Education Association is to encourage ". . . the collection and dissemination of . . . statistical data and other appropriate information relating to rural education" (REA, 1980). A National Rural Education Research Agenda endorsed by the Rural Education Association calls for research relevant to rural education in nine broad categories (Barker and Stephens, 1985).

1. Rural school effectiveness
2. Staff development and professional support
3. Curriculum and instruction
4. Taxonomy of rural education
5. Federal, State, and local policies impacting rural schools and communities
6. Rural school finance
7. School district governance and organization
8. Assessment of rural school assumptions
9. Role of the school in rural development

These themes may not encompass all of the research needs for small/rural schools, but they do establish the major areas in which research is to be focused. Moreover, data collected in these areas will provide policy makers and rural school practitioners information to knowledgeably affect small/rural schools improvement.

In the redesign of the elementary and secondary data program conducted by the National Center for Education Statistics, the Rural Education Association strongly encourages the inclusion of small/rural schools as a specific category in the collection and reporting of data. In light of the various definitions associated with rural education, it would seem that the collection and reporting of data on the basis of school district enrollment size would be the most utilitarian approach. Rural schools have always been, and will likely continue to be, characterized by smallness. According to Sher (1977), small public schools and small school districts have become increasingly rare in America's metropolitan centers. Urban schools and districts have generally always had larger student bodies than rural ones. With continued political and economic pressure to centralize schools, the issues of the small public school will become almost exclusively rural, for rural areas will be the only places such public school exist in significant numbers.

In the collection and reporting of education data, the Rural Education Association recommends that NCES, whenever possible, break down the data based on school districts of fewer than 300 students, 300-999 students, 1000-2500 students, and those in excess of 2500 students. Such a classification would more accurately reflect rural, suburban, and urban similarities and differences. In addition, the availability of comparative data at the national level would provide policy analysts, public educators, and others interested in education with valuable information to assess American public education.

Conclusion

It is impossible to treat rural education as one single or common entity. Rural education encompasses everything from a one-room country schoolhouse in northern Vermont to a sparsely populated western school district responsible for education in a several hundred square mile region. It includes districts having solid financial resources and others with very limited funding sources. Some of America's fastest growing districts, as well as those experiencing the most rapid enrollment decline, are in rural areas. Because of this diversity, much of the effort put forth to improve rural schools can best be met at the local level, where area specific problems can be addressed and treated.

It is not expected that the collection and reporting of national data on the basis of public school district enrollment size will solve the many challenges facing rural educators. Such information will, however, enable local administrators to more knowledgeably assess the operation and management of their own school systems and will provide them with reference information on school systems of similar size. In our nation's quest for excellence in education, the data and information needs for small/rural schools must be included in any collection of statistics conducted by the National Center for Education Statistics.

References

Barker, B. "Research on Texas' Small/Rural Schools." Texas Teacher Education Forum, 10 (1) Spring 1985, pp. 1-8.

Barker, B. and Stephens, R. "A National Rural Education Research Agenda." Report prepared for the Interagency Commission on Rural Education, U. S. Department of Education, May 1985.

CRESS NOTES. "Rural Education and Rural Education Family Education Policy for the 80's." Fact Sheet. ERIC Clearing House on Rural Education and Small Schools. New Mexico State University, Fall 1983/Winter 1984.

Fletcher, J. "The National Seminar on Rural Education." PTA Today, December 1979/January 1980, p.21.

Horn, J. "Higher Education's Response to the Needs of Rural Schools." Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.

Jinks, M. D. "The Advantages of Small High Schools." Rural Education News, 34 (1) Spring 1984, p.3.

Nachtigal, P. Improving Education in Rural Areas: Past Efforts, Future Opportunities. Washington, D.C.: National Institute of Education, December 1979. ERIC Document ED 196-635.

REA News. "Research Focuses on Problems of Rural Schools." Rural Education News, 32 (4), Fall 1984, pp. 4-5.

REA News. "Small School Data will be Separate." Rural Education News, 34 (1) Spring 1983, p. 4.

Rural Education Association, "Goals and Objectives," 1980.

Salmon, P. B. "Small School District Concerns for the 80's." Prepared for the Subcommittee on Elementary, Secondary, and Vocational Education, January 1980. ERIC Document ED 194-502.

Sher, J. (ed) Education in Rural America: A Reassessment of Conventional Wisdom. Boulder, Colorado: Westview Press, 1977.

Special Report on Rural Education. National Advisory Council on the Education of Disadvantaged Children. Washington, D. C.: Government Printing Office, 1979.

Tamblyn, L. R. "The Future of the Small Rural School." Paper prepared for the Montana Educational Association, November 1977.

Treadway, D. Higher Education in Rural America. New York: College Entrance Examination Board, 1984.

U. S. Department of Commerce, Bureau of Census. Characteristics of of the Population. (Vol. 1). Washington, D. C.: Government Printing Office, 1970.

SHEEO/NCES COMMUNICATION NETWORK
Suite 310 • 1860 Lincoln Street • Denver, Colorado 80295 • (303) 830-3687

October 1, 1985

Mr. Emerson Elliott
Administrator
National Center for Education
  Statistics
Brown Building, Room 606
1200 19th Street, N.W.
Washington, D.C. 20208-1404

Dear Mr. Elliott:

I am pleased to provide you with some ideas on the NCES project to redesign its elementary and secondary education data program. This is an important NCES initiative. We applaud the Center's effort to solicit advice from various persons and associations on how NCES might improve these data. The comments which follow were prepared by the State Higher Education Executive Officers/National Center for Education Statistics Communication Network Advisory Committee. I transmit them to you on behalf of the State Higher Education Executive Officers Association.

Most SHEEO agencies use national postsecondary education statistics available from NCES in a variety of ways. By comparison, their use of national elementary/secondary education statistics has been limited. Historically, this has been so because much of the elementary/secondary data used by the agencies has been obtained from their respective state departments of education. Many agencies have not taken advantage of relating national elementary/secondary data to national postsecondary data. Recently, however, the emphasis upon education reform at all levels within the states has generated the need to address interrelated elementary, secondary and postsecondary issues and problems. There are, therefore, some important general comments we feel need to be considered by NCES as the Center works toward redesigning and improving its national elementary/secondary data collection program.

Many education policy issues in the states and at the national level bridge elementary, secondary and postsecondary education. NCES data collections from all sources and levels of education should be compatible so the totality of the educational enterprise and its continuity can be reflected in the data analysis. We need the capacity to assess changes in the educational process at transitional points along the education continuum. The same definitions should be applied to NCES student and institution based survey data so students can be tracked from one level of education to the next. Analysts, as a result, should be able to reliably generalize sampled data to the population (i.e. the sampled data from NLS-88 can be generalized for students in the IPEDS universe of institutions). Elementary/secondary data collections need to be compatible and linked to the NCES Integrated Postsecondary Education Data System (IPEDS) surveys to make possible the ability to interrelate the data. By interrelating the data bases, NCES will be able to show more clearly the multi-dimensional character of education, and thereby demonstrate the need to address many educational problems at more than one area of the education spectrum.

It is important that NCES improve upon both what data are being collected, and upon how the data are collected. Thus, another important element of this redesign effort is the use of current technologies for collecting, transmitting, and disseminating the data to be collected. The timeliness, quality and utility of the NCES data for researchers, administrators, state and national policy makers can be improved if the use of technology to collect and disseminate data is an integral part of the total redesign initiative.

Beyond these general comments, there are some specific data elements, and related information that are needed by postsecondary analysts which should be a part of the NCES elementary/secondary education data collection program. The availability of such data will be of assistance to postsecondary education policy planning and development at the state and national levels.

1. Enrollments in public, private and specialized (state schools for the deaf and blind, etc.), high schools, and enrollments in school-sponsored, home-study programs. The enrollment needed annually from the universe of schools include students in:

    a. Grades 7, 8, 9, 10, 11, 12 by sex, age and ethnicity;

    b. College preparatory, vocational and general curricular tracks for twelfth graders by sex and ethnicity;

    c. Joint high school/collegiate level programs;

    d. Third and fourth year English, mathematics, science and foreign language courses for grade levels 11 and 12;

    e. Different kinds of remedial courses and programs at the secondary and postsecondary levels by sex and ethnicity.

2. Number of high school graduates from public, private and specialized high schools. The number of high school diploma recipients need to be available annually from the universe of schools by sex, age and ethnicity.

3. A survey regarding elementary and high school personnel for the purpose of determining:

   a. Analysis of staff turnover in terms of "quit rates";

   b. Percentage of high school teachers teaching out-of-field;

   c. Measures of teacher quality (i.e. experience, degrees held, test scores, self-reported grade point average in college, etc.)

4. Other information to be provided through data analysis (that perhaps can be derived from existing data) include:

   a. The participation rates of students from different types of secondary schools enrolling in different types of postsecondary institutions;

   b. An analysis of the secondary school courses taken by former high school students enrolling in different types of postsecondary education institutions, and the relationship of such course work to previous academic achievement and test scores in elementary schools.

   c. Basic indicators of the progress being made in improving the condition of education at all levels.

5. Analyses of student-based longitudinal studies that need to be continued include:

   a. High school drop-out rates from all kinds of schools, including private and specialized, by grade level, sex and ethnicity;

   b. Average achievement test scores for students by type of high school (public, private, specialized), location of school (rural, suburban, city), curricular track (college preparation, vocational, general), by student sex and ethnicity; and

   c. Intentions of high school seniors regarding work, military, or education upon completion of their secondary schooling with a six-month follow-up to determine the extent intentions are valid predictors of actual decisions made.

School-based survey data, similarly to postsecondary institution-based data, provide essential information for particular purposes. Obviously, NCES should continue to collect such data. In Addition, student-based survey data are becoming increasingly important. Many policy issues related to drop-outs, remediation, and student course work and achievement can only be addressed through the NCES-sponsored, student-based, longitudinal studies. The information gathered through these studies may, in the future, be the most important elementary/secondary education data series NCES sponsors.

At its annual meeting in July 1985, SHEEO adopted several recommendations it received from its SHEEO/NCES Network Representatives following their national meeting in June (see attachment). Several recommendations relate to improving the relevance, technical quality and utility of NCES data programs to better serve education policy makers at all levels. Implementation of these recommendations will improve NCES' elementary/secondary and postsecondary education data collections. Recommendations 1, 5, 6, 14, 15 and 16 should be considered when redesigning NCES' elementary and secondary education data collections.

On behalf of the SHEEO Association, I thank you for this opportunity to provide our thoughts on the elementary/secondary redesign project. Please contact me for further elaboration on these comments if needed.

Sincerely,

John R. Wittstruck

John R. Wittstruck, Director
SHEEO/NCES Communication Network

JRW:as

cc: Kenneth Ashworth, SHEEO President
SHEEO/NCES Network Advisory Committee
Leslie J. Silverman, NCES
Richard C. Taeuber, NCES

U.S. EQUAL EMPLOYMENT OPPORTUNITY COMMISSION
Washington, D.C. 20507

Mr. Leslie J. Silverman
Deputy Assistant Administrator
Division of Statistical Services
National Center for Education Statistics
Department of Education
Brown Building #413A
1200 19th Street, N.W.
Washington, D.C.

Dear Mr. Silverman:

The U.S. Equal Employment Opportunity Commission is pleased to participate in the National Center for Education Statistics' effort to re-evaluate and redesign its data collection system. We applaud your long-term plans to improve your responsiveness to the needs of various users by providing "cross-sectional and longitudinal data relevant to policy issues and administrative needs, as well as to measurement of our Nation's education systems."

Our input to the project consists of two parts. First, with regard to the kind of data that NCES might consider collecting, we are providing the following general and specific recommendations whose thrust is toward the increased use of case studies and ethnographic research:

1. Collecting data at elementary and secondary school levels indicating actual enrollees by race and ethnic categories;

2. Developing qualitative indicators of primary and elementary school preparation showing self-perception, motivation to learn, and orientation toward school and the world outside of family and neighborhood;

3. Providing interpretive analyses of *Digest of Education Statistics* tables and data sets;

4. Collecting data at both elementary and secondary school levels on drop-out rates by race and ethnic categories. Also providing case studies of successful retention programs, incidences of high drop-out rates, and number of drop-outs who return and complete their studies;

5. Ethnographic studies, particularly in urban areas, on how minority students "move through" the system, with emphasis on barriers and "tracking;"

6. Case studies of the correlation between levels of education and employment by race and ethnic group, in terms of income levels, unemployment and underemployment;
7. Collecting data on students "tracked" during elementary and secondary school as enrollees in general education curriculum, vocational education curriculum, or college preparatory curriculum;
8. Case studies of "more effective" and "less effective" school systems and programs;
9. Case studies comparing public and private schools -- in terms of enrollments by race and ethnic group, quality indicators, objectives, value systems, teaching techniques, administration, and school organization.

Second, in response to your invitation to submit papers addressing select education topics, we are including with this letter two separate papers that discuss the relationship between equal employment opportunity and equal educational opportunity. More specifically, the papers address the critical issues of family, education and employment, with particular emphasis on minority communities. These essays, as well as a third paper that will be sent to you next week, serve as the foundations on which the above recommendations rest.

We ask that you review these papers as "drafts" and not consider these as reflective of official Commission policy. Our intention is to provide research papers for the purpose of stimulating public debate on the nexus between employment and education.

For any further information regarding any aspect of our package, please contact the respective authors or Mark Wong at 634-6750.

Sincerely,

Paul Royston

J. Paul Royston
Director
Office of Program Research

Enclosures

NATIONAL SCIENCE FOUNDATION
WASHINGTON, D.C. 20550

Directorate for Science
and Engineering Education

June 19, 1985

Mr. Leslie J. Silverman
Deputy Assistant Administrator
Division of Statistical Services
National Center for Education Statistics
1200 19th Street, N.W.
Washington, D.C. 20208-1401

Dear Les:

I regret that I was not able to attend the May 28th meeting to discuss the redesign of NCES' elementary and secondary education data program. In my absence, Iris Rotberg represented NSF at that meeting.

In response to your request for papers, NSF is sponsoring two projects at the National Academy of Sciences/National Research Council (NRC) that are of direct relevance to your review. One of these studies represents a year-long effort by the Committee on Indicators of Precollege Science and Mathematics Education, which was charged with proposing a framework for an efficient set of indicators, filling in the framework to the extent possible with existing data, and suggesting data and data analyses that will be needed in the future for a continuing portrayal of the condition of precollege science and mathematics education. NSF is presently supporting a successor committee under the chairmanship of John G. Truxall of the State University at Stony Brook, which is addressing the important goal of developing imaginative new indicators. Jay Noel of NCES has been attending the meetings of the successor committee and has a copy of the initial report.

The second project concerns the supply, demand, and qualifications of teachers of science and mathematics. Under NSF sponsorship, the Committee on National Statistics held a conference on August 9-10, 1984 to identify problems with the available data, or gaps in the data, to discuss problems and possible improvements in the models now used for estimating and projecting supply and demand, and to suggest activities for a follow-on study. Frank Corrigan of NCES has a copy of the conference report and, as you may know, NCES is presently in the process of transferring funds to NSF to support part of a new effort that will be based on recommendations from the August 9-10 conference.

I am enclosing copies of the introductory sections of both of these reports that you may wish to use in addition to this letter, as official submissions for public comment. Of course, the NRC activities as they progress will probably have important implications for the NCES redesign efforts.

In terms of specific comments about NSF's present needs from NCES' elementary and secondary data program, I have the following comments:

1. NCES data on teacher supply and demand and course offerings and enrollment are not broken down in sufficient detail to be useful to NSF for planning purposes. The notable exception is the 1982 transcript data from High School and Beyond, which produced enrollment data on detailed of science courses (such as physics, chemistry, and earth sciences) and mathematics courses (such as algebra I, algebra II, geometry, trigonometry, and calculus).

2. Time series data on course enrollment in science and mathematics disciplines are largely lacking. NCES should have more consistency in the design of surveys, data collection, and analysis. Also, more and better data are needed on the amount of time students spend on homework.

3. The most significant determinant of teacher demand projections are turnover rates (which appear to be age specific). Yet NCES data on teacher turnover rates are several years out of date and even these earlier data are not age specific. Again, NCES projections of teacher demand are not broken down by science and mathematics disciplines.

4. Supply projections are largely dependent on new teacher graduates. NCES uses the questionable practice of projecting new teacher graduates based on estimates of percentages of total bachelors degrees granted and these estimates are aggregated so that data on science and mathematics disciplines are not available.

5. Almost no data are available on the reserve pool of teachers and the number who return to teach.

6. The NCES practice of counting teacher vacancies leaves a lot to be desired in trying to determine the extent of shortages of qualified teachers of science and mathematics. For example, there is evidence that many science and mathematics teachers are teaching out-of-field because of shortages. Also, it is not clear whether a teacher certified in both mathematics and chemistry would be counted as a mathematics or a chemistry teacher, or both.

7. Adequate information is lacking on the qualifications of teachers who are responsible for teaching science and mathematics in high school, middle school, and elementary school. In many instances, certification is not a good proxy for teacher qualifications because of disparate certification practices of states. We also need to have more data on these state certification practices.

8. In terms of curriculum content, periodic surveys should be conducted of use of various science and mathematics textbooks at each grade level. Surveys of textbook use should be followed by content analysis of the most frequently used textbooks.

I appreciate the high degree of cooperation that you and your collegues at NCES have had with NSF in the recent past and I look forward to working with you on these issues in the future.

Sincerely yours,

Dick

Richard Berry
Program Director
Studies and Analyses

Enclosures

cc: W. Gillespie, SEE
I. Rotberg, SEE/OSPA

THE OFFICE OF THE ASSISTANT SECRETARY OF DEFENSE
WASHINGTON, D.C. 20301-4000

FORCE MANAGEMENT
AND PERSONNEL
(Military Personnel & Force Management)

1 AUG 1985

Mr. Leslie J. Silverman
National Center for Education Statistics
1200 19th Street, N.W.
Washington, D.C. 20208-1401

Dear Mr. Silverman:

As we discussed at our July 19, 1985 meeting, the Department of Defense would be pleased to participate in the Elementary/Secondary Education Data Program Redesign Project and we have specific data we would like you to consider collecting on our behalf.

Our secondary school data collection request, outlined on the attachment, stems from our need for current, consistent and reliable data in support of military recruiting. Specifically, a key portion of the Recruit Market Network (RMN), a major DoD data base available to recruiters through a nationwide teleprocessing network, includes information about the high school population. Because the high school data for inclusion in the RMN is derived largely from recruiters, private firms, and secondary sources, we would be delighted to replace these sources with your data.

If you have questions about the attachment or plan to host future meetings on this topic, we would appreciate your contacting Zahava D. Doering, Chief, Survey and Market Analysis Division, Defense Manpower Data Center, at 696-5826. Dr. Doering, or members of her staff, will attend relevant meetings and can provide additional information.

Sincerely,

Anita R. Lancaster

Anita R. Lancaster
Assistant Director
Accession Policy

Attachment

HIGH SCHOOL INFORMATION AND DATA TO SUPPORT MILITARY RECRUITING:

REQUIREMENTS

Background

o The Recruit Market Network (RMN) is a common data base, available to users (on-line) through a teleprocessing network, established to support recruiting efforts in the Military Services.

o A key portion of the RMN is devoted to information about the nation's high school population, organized in ways to aid decision makers in allocating their resources.

o The cataloguing, tabulating and associated problems of determining school locations, current enrollment characteristics, and public and private school inventories has remained an on-going problems for recruiters, managers, and DoD officials.

o Data currently available from the National Center for Education Statistics (NCES) does not provide the required detailed data.

Objective

o Provide the recruiting community, through the RMN, with a reliable, up-to-date, efficient, data base containing information on the nation's high schools.

o Eliminate the need to utilize recruiters, private firms, and secondary data sources for this information.

Requirements (each school)

o The RMN requires a data file with a record for every high school which provides the following information:

- Type of School

  - Public
  - Catholic
  - Private (Non-Catholic)
  - Vocational

- School Location

  - City
  - County
  - State
  - Zip-Code
  - School District

- Senior Class enrollment counts, by sex and race/ethnic
- Junior Class enrollment counts, by sex and race/ethnic
- Previous Year's graduates counts, by sex and race/ethnic
- Percentage attending college (for previous years' graduates)

o These data elements are needed on an annual basis